

192663

**STIC-Biotech/ChemLib**

**From:** Ramirez, Delia  
**Sent:** Monday, June 12, 2006 3:12 PM  
**To:** STIC-Biotech/ChemLib  
**Subject:** 10/612779

Hi,

I would like to request the following searches: SEQ ID NO:6 and 30 in the protein databases (commercial & interference).

Please provide a hard copy of this search.

Thank you very much,

Delia M. Ramirez, Ph.D.  
Patent Examiner  
Recombinant Enzymes-Art Unit 1652  
USPTO  
400 Dulany Street, Remsen Bldg., 2D74, Mail room 2C70  
Alexandria, VA 22314  
(571) 272-0938  
delia.ramirez@uspto.gov

78201

CRFE

**BEST AVAILABLE COPY**

\*\*\*\*\*

Searcher: \_\_\_\_\_  
Searcher Phone: \_\_\_\_\_  
Date Searcher Picked up: \_\_\_\_\_  
Date completed: \_\_\_\_\_  
Searcher Prep Time: \_\_\_\_\_  
Online Time: \_\_\_\_\_

\*\*\*\*\*

Type of Search  
NA# \_\_\_\_\_ AA# \_\_\_\_\_  
S/L: \_\_\_\_\_ Oligomer: \_\_\_\_\_  
Encode/Transl: \_\_\_\_\_  
Structure #: \_\_\_\_\_ Text: \_\_\_\_\_  
Inventor: \_\_\_\_\_ Litigation: \_\_\_\_\_

\*\*\*\*\*

Vendors and cost where applicable  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
QUESTEL/ORBIT: \_\_\_\_\_  
LEXIS/NEXIS: \_\_\_\_\_  
SEQUENCE SYSTEM: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (Specify): \_\_\_\_\_

**THIS PAGE IS BLANK**

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 14, 2006, 15:14:44 ; Search time 286.262 Seconds  
(without alignments)  
972.693 Million cell updates/sec

Title: US-10-612-779-6

Perfect score: 3089

Sequence: 1 MCGIVGATAQRDVAREILLEG.....LIKGTVDVQPRNLAKSVTIVE 609

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq 8:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*
- 9: Geneseqp2005s:\*
- 10: Geneseqp2006s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3089	100.0	609	3 AAY58824	Aay58824 E. coli g
2	3089	100.0	609	8 ADI38861	Adi38861 Mutant gl
3	3089	100.0	609	8 ADU00592	Adu00592 Amino aci
4	3072	99.4	609	3 AAY58827	Aay58827 E. coli g
5	3072	99.4	609	3 AAY58826	Aay58826 E. coli g
6	3072	99.4	609	8 ADI38865	Adi38865 Mutant gl
7	3072	99.4	609	8 ADI38867	Adi38867 Mutant gl
8	3068	99.3	609	3 AAY58822	Aay58822 E. coli g
9	3068	99.3	609	4 AAU34806	Aau34806 E. coli c
10	3068	99.3	609	6 ABU28829	Abu28829 Protein e
11	3068	99.3	609	8 ADI38857	Adi38857 Glucosami
12	3068	99.3	609	8 ADS45181	Ads45181 Bacterial
13	3068	99.3	609	8 ADU00590	Adu00590 Amino aci
14	3063	99.2	608	9 ADW23840	Adw23840 Novel hum
15	3061	99.1	609	3 AAY58825	Aay58825 E. coli g
16	3061	99.1	609	3 AAY58863	Adi38863 Mutant gl
17	3053	98.8	609	3 AAY58823	Aay58823 E. coli g
18	3053	98.8	609	8 ADI38859	Adi38859 Mutant gl
19	3043	98.5	609	8 ADI38869	Adi38869 Mutant gl
20	3035	98.3	609	4 AAU38472	Aau38472 Salmonell
21	3035	98.3	609	6 ABU48228	Abu48228 Protein e
22	2932	94.9	609	6 ABU28355	Abu28355 Protein e
23	2862	92.7	609	4 AAU36165	Aau36165 Klebsiell

24	2862	92.7	651	7 ABO64174	ABO64174 Klebsiell
25	2857	92.5	609	6 ABU31411	ABU31411 Protein e
26	2700.5	87.4	608	6 ABU45041	ABU45041 Protein e
27	2668	86.4	609	6 ABU50609	ABU50609 Protein e
28	2647	85.7	609	8 ADS42812	Ads42812 Bacterial
29	2605	84.3	609	8 ADN17785	Adn17785 Bacterial
30	2558	82.8	610	6 ABM70565	Abm70565 Phototrab
31	2498.5	80.9	622	7 ADF07237	Adf07237 Bacterial
32	2493.5	80.7	608	6 ABU41104	ABU41104 Protein e
33	2256.5	73.0	610	4 AAU35425	Aau35425 Haemophil
34	2256.5	73.0	610	6 ABU30239	ABU30239 Protein e
35	2256.5	73.0	610	6 AAE30459	Aae30459 Haemophil
36	2256.5	73.0	610	10 AEE31344	Aee31344 Haemophil
37	2250.5	72.9	610	6 ABU49139	ABU49139 Protein e
38	2188.5	70.8	610	6 ABU39464	ABU39464 Protein e
39	1993	64.5	611	6 ABU39730	ABU39730 Protein e
40	1990.5	64.4	610	8 ADS24814	Ads24814 Bacterial
41	1983	64.2	611	4 AAU36518	Aau36518 Pseudomon
42	1983	64.2	611	6 ABU38924	ABU38924 Protein e
43	1983	64.2	616	7 ABO70418	ABO70418 Pseudomon
44	1975	63.9	611	6 ABU41714	ABU41714 Protein e
45	1797.5	58.2	612	6 ABU17019	ABU17019 Protein e

#### ALIGNMENTS

RESULT 1	
AAY58824	
ID	AAY58824 standard; protein; 609 AA.
XX	
AC	AAY58824;
XX	
DT	08-MAY-2000 (first entry)
XX	
DE	E. coli glucosamine-6-phosphate synthase mutant GlcN6P-S-54.
XX	
KW	Glucosamine-6-phosphate synthase; glmS gene; mutant; GlcN6P-S-54;
KW	glucosamine; metabolic engineering; plasmid pKLN23-54; mutcin.
XX	
OS	Escherichia coli.
XX	
FH	Key Location/Qualifiers
FT	Misc-difference 39 /note= "replaces wild-type Ala"
FT	Misc-difference 250 /note= "replaces wild-type Arg"
FT	Misc-difference 472 /note= "replaces wild-type Gly"
FT	
PN	WO200004182-A1.
XX	
PD	27-JAN-2000.
XX	
PF	15-JUL-1999; 99WO-US015976.
XX	
PR	15-JUL-1998; 98US-00115475.
XX	
PA	(DCVB-) DCV INC DBA BIO-TECH RESOURCES.
XX	
PI	Berry A, Burlingame RP, Millis JR;
XX	
DR	WPI; 2000-182441/16.
DR	N-PSDB; AAZ58251.
XX	
PT	Fermentation of E. coli having an altered amino acid sugar metabolic pathway to produce glucosamine, especially using novel recombinant variant glucosamine-6-phosphate synthases.
XX	
PS	Claim 28; Page 124-126; 150pp; English.
XX	
CC	The present sequence is that of a mutant, denoted GlcN6P-S-54, of the glucosamine-6-phosphate synthase (GlcN6P synthase) of Escherichia coli.
CC	

CC When compared with the wild-type sequence (see AAY58822), the mutant  
CC includes A39T, R250C and G472S amino acid substitutions. These  
CC alterations are predicted from the mutated glms gene in plasmid pKLN23-54  
CC (see AAZ58251). The invention provides methods for the overproduction of  
CC glucosamine by fermentation using a genetically engineered microorganism,  
CC especially E. coli, that includes a modified GlcN6P synthase. Production  
CC of the glucosamine by recombinant strain 2123-54 (pKLN23-54) was  
CC significantly increased when compared to a strain expressing wild-type  
CC Glc6NP synthase owing to reduced product inhibition. Glucosamine  
CC concentrations of over 12 g/l have been obtained  
XX  
XX

SQ Sequence 609 AA;

Query Match 100.0%; Score 3089; DB 3; Length 609;  
Best Local Similarity 100.0%; Pred. No. 2.7e-271;  
Matches 609; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCGIVGATAQRDVAEILLEGRLRLEYRGYDSAGLAVVDTEGHMTLRLKGVQMLAQAAE 60  
DB 1 MCGIVGATAQRDVAEILLEGRLRLEYRGYDSAGLAVVDTEGHMTLRLKGVQMLAQAAE 60  
QY 61 EHPHGGTGIATRWATHGEPSEVNAHPVSEHIVVHNGIIEHPELREELKARGYTFV 120  
DB 61 EHPHGGTGIATRWATHGEPSEVNAHPVSEHIVVHNGIIEHPELREELKARGYTFV 120  
QY 121 SETDTEVIAHLVNNELKOGGTLREAVLRAIPQLRGAYGTIVMSRHPDTLLAARSGSPV 180  
DB 121 SETDTEVIAHLVNNELKOGGTLREAVLRAIPQLRGAYGTIVMSRHPDTLLAARSGSPV 180  
QY 181 IGLGMGENFIASDQALLPVTRRFIFLEEGDIAETIRRSVNI FDKTGAEVKRQDIESNLQ 240  
DB 181 IGLGMGENFIASDQALLPVTRRFIFLEEGDIAETIRRSVNI FDKTGAEVKRQDIESNLQ 240  
QY 241 YDAGDKGIYCHYMOKEIYEQPNAIKNTLTGRISHGQVDLSBELGPNADLLSKVEHIQILA 300  
DB 241 YDAGDKGIYCHYMOKEIYEQPNAIKNTLTGRISHGQVDLSBELGPNADLLSKVEHIQILA 300  
QY 301 CGTSYNSGMVSRYPWFESLAGIPCDVEIASPEFRYKSAVRRNSLMTLSQSGETADTLAGL 360  
DB 301 CGTSYNSGMVSRYPWFESLAGIPCDVEIASPEFRYKSAVRRNSLMTLSQSGETADTLAGL 360  
QY 361 RLSKELGVLGSLAICNVPSSIVRESDLALMTNAGTEIGVASTKFTTQLTVLLMLVAKL 420  
DB 361 RLSKELGVLGSLAICNVPSSIVRESDLALMTNAGTEIGVASTKFTTQLTVLLMLVAKL 420  
QY 421 SRLKGLDASIEHDIHVHGLQALPSRIEQMSQDKRIEALAEFSDKHAFILSRGDQYPIA 480  
DB 421 SRLKGLDASIEHDIHVHGLQALPSRIEQMSQDKRIEALAEFSDKHAFILSRGDQYPIA 480  
QY 481 LEGALKLKEISYTHAEVAAGELKHGPLALIDAMPVIVVAPNNELLEKLKSNTEEVRRAR 540  
DB 481 LEGALKLKEISYTHAEVAAGELKHGPLALIDAMPVIVVAPNNELLEKLKSNTEEVRRAR 540  
QY 541 GGQLYVFADQAGFVSSDNMHIEMPHVEEVIAPFTYVLPQLLAYHVALLKGTDDVQPR 600  
DB 541 GGQLYVFADQAGFVSSDNMHIEMPHVEEVIAPFTYVLPQLLAYHVALLKGTDDVQPR 600  
QY 601 NLAKSVTVE 609  
DB 601 NLAKSVTVE 609

RESULT 2  
ADI38861  
ADI38861 standard; protein; 609 AA.

XX ADI38861;

DT 15-APR-2004 (first entry)

DE Mutant glucosamine-6-phosphate synthase, glms\*54, SEQ ID 6.

XX Glucosamine; N-acetylglucosamine; fermentation;

glucosamine-6-phosphate acetyltransferase;  
glucosamine-6-phosphate synthase; glucosamine-6-phosphate deaminase;  
glucosamine-1-phosphate N-acetyltransferase; glucosamine-6-phosphate;  
glucosamine-1-phosphate; N-acetylglucosamine-1-phosphate;  
N-acetylglucosamine-6-phosphate; enzyme.

OS Escherichia coli.

OS Synthetic.

PN WO2004003175-A2.

XX 08-JAN-2004.

XX 01-JUL-2003; 2003WO-US020925.

XX 01-JUL-2002; 2002US-0393348P.

XX (ARKI-) ARKION LIFE SCI LLC.

PI Deng M, Angerer JD, Cyron D, Grund AD, Jerrell TA, Leanna C;  
PI Mathre O, Rosson R, Running J, Severson D, Song L, Wassink S;

XX WPI: 2004-203380/19.

XX N-PSDB; ADI38860.

Producing glucosamine or N-acetylglucosamine by fermentation involves  
culturing microorganism comprising glucosamine-6-phosphate  
acetyltransferase, in fermentation medium, and collecting product.

Claim 15; SEQ ID NO 6; 327pp; English.

The present invention relates to a method (M1) for producing glucosamine  
and N-acetylglucosamine by fermentation. The method comprises (a)  
culturing in a fermentation medium a microorganism (I) which comprises  
endogenous glucosamine-6-phosphate acetyltransferase (II) and a genetic  
modification that increases the activity of (II), glucosamine-6-phosphate  
synthase (III) or glucosamine-6-phosphate deaminase (IV), or decreases  
the activity of (IV) and increases the activity of glucosamine-1-  
phosphate N-acetyltransferase (V), and (b) and collecting the product,  
which is chosen from the group consisting of glucosamine-6-phosphate,  
glucosamine, glucosamine-1-phosphate, N-acetylglucosamine-1-phosphate, N-  
acetylglucosamine-6-phosphate, and N-acetylglucosamine. The present  
sequence was used to illustrate the method of the invention.

SQ Sequence 609 AA;

Query Match 100.0%; Score 3089; DB 8; Length 609;  
Best Local Similarity 100.0%; Pred. No. 2.7e-271;  
Matches 609; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCGIVGATAQRDVAEILLEGRLRLEYRGYDSAGLAVVDTEGHMTLRLKGVQMLAQAAE 60  
DB 1 MCGIVGATAQRDVAEILLEGRLRLEYRGYDSAGLAVVDTEGHMTLRLKGVQMLAQAAE 60  
QY 61 EHPHGGTGIATRWATHGEPSEVNAHPVSEHIVVHNGIIEHPELREELKARGYTFV 120  
DB 61 EHPHGGTGIATRWATHGEPSEVNAHPVSEHIVVHNGIIEHPELREELKARGYTFV 120  
QY 121 SETDTEVIAHLVNNELKOGGTLREAVLRAIPQLRGAYGTIVMSRHPDTLLAARSGSPV 180  
DB 121 SETDTEVIAHLVNNELKOGGTLREAVLRAIPQLRGAYGTIVMSRHPDTLLAARSGSPV 180  
QY 181 IGLGMGENFIASDQALLPVTRRFIFLEEGDIAETIRRSVNI FDKTGAEVKRQDIESNLQ 240  
DB 181 IGLGMGENFIASDQALLPVTRRFIFLEEGDIAETIRRSVNI FDKTGAEVKRQDIESNLQ 240  
QY 241 YDAGDKGIYCHYMOKEIYEQPNAIKNTLTGRISHGQVDLSBELGPNADLLSKVEHIQILA 300  
DB 241 YDAGDKGIYCHYMOKEIYEQPNAIKNTLTGRISHGQVDLSBELGPNADLLSKVEHIQILA 300  
QY 301 CGTSYNSGMVSRYPWFESLAGIPCDVEIASPEFRYKSAVRRNSLMTLSQSGETADTLAGL 360  
DB 301 CGTSYNSGMVSRYPWFESLAGIPCDVEIASPEFRYKSAVRRNSLMTLSQSGETADTLAGL 360



QY 361 RLSKELGYLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTTLTVLLMLVAKL 420  
 Db 361 RLSKELGYLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTTLTVLLMLVAKL 420  
 QY 421 SRLKGLDASIEHDIHVHGLQALPSRIEQMLSQDKRIEALAEFSDKHHLFLSRGQOYPIA 480  
 Db 421 SRLKGLDASIEHDIHVHGLQALPSRIEQMLSQDKRIEALAEFSDKHHLFLSRGQOYPIA 480  
 QY 481 LEGALKLKEISYTHAEYAAAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEBVRAR 540  
 Db 481 LEGALKLKEISYTHAEYAAAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEBVRAR 540  
 QY 541 GGQLYVFADQDAGFVSSDNNHIIEMPHVEEVIAPFYTVPLQLLAVHVALIKGTDVQDR 600  
 Db 541 GGQLYVFADQDAGFVSSDNNHIIEMPHVEEVIAPFYTVPLQLLAVHVALIKGTDVQDR 600  
 QY 601 NLAKSVTVE 609  
 Db 601 NLAKSVTVE 609

RESULT 3  
 ADU00592  
 ID ADU00592 standard; protein; 609 AA.  
 AC ADU00592;  
 XX  
 XX 13-JAN-2005 (first entry)  
 XX  
 DE Amino acid sequence of a mutant glucosamine-6-phosphate synthase.  
 KW chitin; chitosan; fermentation;  
 KW glucamine-fructose-6-phosphate amidotransferase; glucosamine synthetase;  
 KW glucosamine-6-phosphate synthase; GFA1 gene; glms gene;  
 KW glucosamine-6-phosphate acetyltransferase; GNA1 gene; chitin synthase;  
 KW chitin deacetylase; CDA1 gene; CDA2 gene;  
 KW N-acetylglucosamine-6-phosphate deacetylase;  
 KW glucosamine-6-phosphate deaminase; chitinase; chitosanase; fungus; yeast;  
 KW enzyme; glms\*54.  
 XX  
 OS Escherichia coli.  
 XX  
 PN WO2004092391-A2.  
 XX  
 PD 28-OCT-2004.  
 XX  
 PF 12-APR-2004; 2004WO-US011286.  
 XX  
 PR 11-APR-2003; 2003US-0462087P.  
 XX  
 PA (ARKI-) ARKION LIFE SCI LLC.  
 XX  
 PI Deng M, McMullin TW, Grund AD;  
 XX  
 DR WPI: 2004-766980/75.  
 DR N-PSDB; ADU00591.  
 XX  
 PT Producing chitin or chitosan, involves culturing microorganism with  
 PT genetic modification that results in increase in activity of glutamine-  
 PT fructos-6-phosphate amidotransferase, in fermentation medium, to produce  
 PT chitin or chitosan.  
 XX

Example 1; SEQ ID NO 25; 161pp; English.

CC The specification describes a method for producing chitin or chitosan by  
 CC a fermentation process. The method involves culturing in a fermentation  
 CC medium a microorganism which comprises one or more genetic modifications  
 CC that result in an increase in the activity of glutamine-fructose-6-  
 CC phosphate amidotransferase (also known as glucosamine synthetase and  
 CC glucosamine-6-phosphate synthase, and encoded by the GFA1 eukaryotic gene  
 CC and the glms bacterial gene), glucosamine-6-phosphate acetyltransferase  
 CC (encoded by GNA1), chitin synthase or chitin deacetylase (encoded by CDA1

CC and CDA2), or in decrease in the activity of N-acetylglucosamine-6-  
 CC phosphate deacetylase, glucosamine-6-phosphate deaminase, chitinase and  
 CC chitosanase, and collecting the chitin or chitosan. The method is useful  
 CC for producing chitin or chitosan by utilizing microorganisms such as  
 CC fungus, yeast (e.g. Saccharomyces or Schizosaccharomyces) and  
 CC filamentous fungus (e.g. Aspergillus, Absidia or Rhizopus), preferably S.  
 CC cerevisiae, A. niger or A. nidulans. The method enables high quantities  
 CC of chitin and chitosan to be produced cost effectively. The present  
 CC sequence represents a mutant glucosamine-6-phosphate synthase, encoded by  
 CC glms\*54. This enzyme is resistant to feedback inhibition by glucosamine-6-  
 CC phosphate. The polynucleotide was used to transform yeast for use in the  
 XX method of the invention.

SQ Sequence 609 AA;

Query Match 100.0%; Score 3089; DB 8; Length 609;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-271;  
 Matches 609; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCGIVGAIAQRDVAEILLEGRLRLEYRGYDVSAGLAVDTEGHMTLRLRIGKVKQMLAQAAE 60  
 Db 1 MCGIVGAIAQRDVAEILLEGRLRLEYRGYDVSAGLAVDTEGHMTLRLRIGKVKQMLAQAAE 60  
 QY 61 EHLPGGTGIAHTRWATHGEPSEVNAHPVSHIIVVHNGIIEENHEPLREELKARGYTFV 120  
 Db 61 EHLPGGTGIAHTRWATHGEPSEVNAHPVSHIIVVHNGIIEENHEPLREELKARGYTFV 120  
 QY 121 SETDTEVIHLVNWELKOGGTLREAVLRAIPOLRGAYGTIVMDSRHPDTLLAARSGSPLV 180  
 Db 121 SETDTEVIHLVNWELKOGGTLREAVLRAIPOLRGAYGTIVMDSRHPDTLLAARSGSPLV 180  
 QY 181 IGLNGENFIASDQALLPVTRRFIFLEEGDIAETIRRSVNIFFDKTGARVKQDIESNLQ 240  
 Db 181 IGLNGENFIASDQALLPVTRRFIFLEEGDIAETIRRSVNIFFDKTGARVKQDIESNLQ 240  
 QY 241 YDAGDKGIYCHYMQKEIYEQPNAIKNTLTGRIHQGVLDSELGPNADLLSKVEHIQILA 300  
 Db 241 YDAGDKGIYCHYMQKEIYEQPNAIKNTLTGRIHQGVLDSELGPNADLLSKVEHIQILA 300  
 QY 301 CGTSYNSGMVSRYPFESLAGIPCDVEIASFEPRYKSAVRNLSMITLSQSGETADTLA 360  
 Db 301 CGTSYNSGMVSRYPFESLAGIPCDVEIASFEPRYKSAVRNLSMITLSQSGETADTLA 360  
 QY 361 RLSKELGYLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTTLTVLLMLVAKL 420  
 Db 361 RLSKELGYLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTTLTVLLMLVAKL 420  
 QY 421 SRLKGLDASIEHDIHVHGLQALPSRIEQMLSQDKRIEALAEFSDKHHLFLSRGQOYPIA 480  
 Db 421 SRLKGLDASIEHDIHVHGLQALPSRIEQMLSQDKRIEALAEFSDKHHLFLSRGQOYPIA 480  
 QY 481 LEGALKLKEISYTHAEYAAAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEBVRAR 540  
 Db 481 LEGALKLKEISYTHAEYAAAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEBVRAR 540  
 QY 541 GGQLYVFADQDAGFVSSDNNHIIEMPHVEEVIAPFYTVPLQLLAVHVALIKGTDVQDR 600  
 Db 541 GGQLYVFADQDAGFVSSDNNHIIEMPHVEEVIAPFYTVPLQLLAVHVALIKGTDVQDR 600  
 QY 601 NLAKSVTVE 609  
 Db 601 NLAKSVTVE 609

RESULT 4

AAV58827  
 ID AAV58827 standard; protein; 609 AA.

AC AAV58827;

DT 08-MAY-2000 (first entry)

DE E. coli glucosamine-6-phosphate synthase mutant GlcN6P-S-151.

```

XX KW Glucosamine-6-phosphate synthase; glms gene; mutant; GlcN6P-S-151;
KW glucosamine; metabolic engineering; plasmid pKLN23-151; mutein.
XX OS Escherichia coli.
XX FH Key Location/Qualifiers
FT Misc-difference 472
FT /note= "replaces wild-type Gly"
XX PN WO200004182-A1.
XX XX
XX PD 27-JAN-2000.
XX PF 15-JUL-1999; 99WO-US015976.
XX PR 15-JUL-1998; 98US-00115475.
XX PA (DCVB-) DCV INC DBA BIO-TECH RESOURCES.
XX PI Berry A, Burlingame RP, Millis JR;
XX WPI; 2000-182441/16.
XX DR N-PSDB; AAZ58254.
XX PT Fermentation of E. coli having an altered amino acid sugar metabolic
PT pathway to produce glucosamine, especially using novel recombinant
PT variant glucosamine-6-phosphate synthases.
XX PS Claim 28; Page 143-145; 150pp; English.
XX CC The present sequence is that of a mutant, denoted GlcN6P-S-151, of the
CC glucosamine-6-phosphate synthase (GlcN6P synthase) of Escherichia coli.
CC When compared with the wild-type sequence (see AAY58822), the mutant
CC includes a Ser for Gly-472 amino acid substitution. This alteration was
CC predicted from the mutated glms gene in plasmid pKLN23-151 (see AAZ58254).
CC The invention provides methods for the overproduction of glucosamine by
CC fermentation using a genetically engineered microorganism, especially E.
CC coli, that includes a modified GlcN6P synthase. Production of the
CC glucosamine by recombinant strain 2123-149 (pKLN23-149) was marginally
CC increased when compared to a strain expressing wild-type GlcN6P synthase
XX Sequence 609 AA;
XX
XX Query Match 99.4%; Score 3072; DB 3; Length 609;
XX Best Local Similarity 99.7%; Pred. No. 9.7e-270;
XX Matches 607; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MCGIVGAIQAORDVAEILLEGRLREYRGYDVSAGLAVVDTEGHMTLRLGKVMQAQAE 60
DB 1 MCGIVGAIQAORDVAEILLEGRLREYRGYDVSAGLAVVDTEGHMTLRLGKVMQAQAE 60
QY 61 EHLPLGGTGIAHTRWATHGESEVNAHPVSEHIVVHNGIIEHPELREELKARGYTFV 120
DB 61 EHLPLGGTGIAHTRWATHGESEVNAHPVSEHIVVHNGIIEHPELREELKARGYTFV 120
QY 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPOLRGAYGTVMDSRPHPTLLAARSGSPV 180
DB 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPOLRGAYGTVMDSRPHPTLLAARSGSPV 180
QY 181 IGLNGENFIASDQALLPVTRRFTFLEEGDIAETTRSVNIFDKTGAEVKRODIESNLQ 240
DB 181 IGLNGENFIASDQALLPVTRRFTFLEEGDIAETTRSVNIFDKTGAEVKRODIESNLQ 240
QY 241 YDAGDKGIYCHYMQKEIYEQPNAINKTITGRISHGQVDLSLSELPNADELLSKVEHIQILA 300
DB 241 YDAGDKGIYCHYMQKEIYEQPNAINKTITGRISHGQVDLSLSELPNADELLSKVEHIQILA 300
QY 301 CGTSSNSGMVSRYPFESLAGIPCDVEIASERFPRYKSAVRNLSLMTITLQSGETADTLA 360
DB 301 CGTSSNSGMVSRYPFESLAGIPCDVEIASERFPRYKSAVRNLSLMTITLQSGETADTLA 360
QY 361 RLSKELGYLGSIAICNVPGSSIVRSDALMTNAGTEIGVASTKRAFTTQLTVLLMLVAKL 420

```

```

DB 361 RLSKELGYLGSIAICNVPGSSIVRSDALMTNAGTEIGVASTKRAFTTQLTVLLMLVAKL 420
QY 421 SRLKGLDASIEHDIVHGLQALPSRIEQWLSQDKRIEALAEFSDKHAFLSRSGDQYPIA 480
DB 421 SRLKGLDASIEHDIVHGLQALPSRIEQWLSQDKRIEALAEFSDKHAFLSRSGDQYPIA 480
QY 481 LEGALKLKEISYTHAEAYAGELKHGPLALIDAMPVIVVAPNNELLEKLKSNIEEVRAR 540
DB 481 LEGALKLKEISYTHAEAYAGELKHGPLALIDAMPVIVVAPNNELLEKLKSNIEEVRAR 540
QY 541 GGQLYVFADQDAGFVSSDNMHIIEMPHVEEVIAPFTVTPQLLAYHVALIKGTDVDDQPR 600
DB 541 GGQLYVFADQDAGFVSSDNMHIIEMPHVEEVIAPFTVTPQLLAYHVALIKGTDVDDQPR 600
QY 601 NLAKSVTVE 609
DB 601 NLAKSVTVE 609

```

```

RESULT 5
AAY58826
ID AAY58826 standard; protein; 609 AA.
XX AC AAY58826;
XX DT 08-MAY-2000 (first entry)
XX DE E. coli glucosamine-6-phosphate synthase mutant GlcN6P-S-149.
XX KW Glucosamine-6-phosphate synthase; glms gene; mutant; GlcN6P-S-149;
XX glucosamine; metabolic engineering; plasmid pKLN23-149; mutein.
XX OS Escherichia coli.
XX FH Key Location/Qualifiers
FT Misc-difference 472 /note= "replaces wild-type Gly"
XX PN WO200004182-A1.
XX PD 27-JAN-2000.
XX PF 15-JUL-1999; 99WO-US015976.
XX PR 15-JUL-1998; 98US-00115475.
XX PA (DCVB-) DCV INC DBA BIO-TECH RESOURCES.
XX PI Berry A, Burlingame RP, Millis JR;
XX WPI; 2000-182441/16.
XX DR N-PSDB; AAZ58253.
XX PT Fermentation of E. coli having an altered amino acid sugar metabolic
PT pathway to produce glucosamine, especially using novel recombinant
PT variant glucosamine-6-phosphate synthases.
XX PS Claim 28; Page 137-139; 150pp; English.
XX CC The present sequence is that of a mutant, denoted GlcN6P-S-149, of the
CC glucosamine-6-phosphate synthase (GlcN6P synthase) of Escherichia coli.
CC When compared with the wild-type sequence (see AAY58822), the mutant
CC includes a Ser for Gly-472 amino acid substitution. This alteration was
CC predicted from the mutated glms gene in plasmid pKLN23-149 (see AAZ58253).
CC The invention provides methods for the overproduction of glucosamine by
CC fermentation using a genetically engineered microorganism, especially E.
CC coli, that includes a modified GlcN6P synthase. Production of the
CC glucosamine by recombinant strain 2123-149 (pKLN23-149) was marginally
CC increased when compared to a strain expressing wild-type GlcN6P synthase
XX Sequence 609 AA;

```

Query Match	99.4%;	Score 3072;	DB 3;	Length 609;
Best Local Similarity	99.7%;	Pred. No. 9.7e-270;		
Matches 607;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;
QY	1	MCGIVGAIQAQRDVAEILLEGRLREYRGYDSAGLAVVDTEGHMTRLRRLRGKQVOMLAQAAE	60	
Db	1	MCGIVGAIQAQRDVAEILLEGRLREYRGYDSAGLAVVDTEGHMTRLRRLRGKQVOMLAQAAE	60	
QY	61	EHPHGGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHPELREELKARGYTFV	120	
Db	61	EHPHGGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHPELREELKARGYTFV	120	
QY	121	SETDTEVIAHLVNWELKOGGTLREAVLRAIPQLRGAYGTVMDSRHPDTLLAARSSPLV	180	
Db	121	SETDTEVIAHLVNWELKOGGTLREAVLRAIPQLRGAYGTVMDSRHPDTLLAARSSPLV	180	
QY	181	IGLGMGENFIASDQALLPVTRRFIFLEEGDIAEITRRSVNIFDKTGAEVKRQDIESNLQ	240	
Db	181	IGLGMGENFIASDQALLPVTRRFIFLEEGDIAEITRRSVNIFDKTGAEVKRQDIESNLQ	240	
QY	241	YDAGDKGIYCHYMQKEIYEOPNAIKNTLTGRISHGQVDLSELGNADLLSKVEHIQILA	300	
Db	241	YDAGDKGIYCHYMQKEIYEOPNAIKNTLTGRISHGQVDLSELGNADLLSKVEHIQILA	300	
QY	301	CGTSYNSGMVSRVWFESLAGIPCDVEIASFPRYKSAVRNSLMTLSQSGETADTLA	360	
Db	301	CGTSYNSGMVSRVWFESLAGIPCDVEIASFPRYKSAVRNSLMTLSQSGETADTLA	360	
QY	361	RLSKELGYLGSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFVTTQTLTLLMLVAKL	420	
Db	361	RLSKELGYLGSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFVTTQTLTLLMLVAKL	420	
QY	421	SRKGLDASIEHDI VHGLQALPSRIEOMLSQDKRIEALAEDFSCKHHALFLSRGDQYPIA	480	
Db	421	SRKGLDASIEHDI VHGLQALPSRIEOMLSQDKRIEALAEDFSCKHHALFLSRGDQYPIA	480	
QY	481	LEGALKLKEISYTHAEYAAAGELKHGFLALIDADMPVIVVAPNNLEKLSKSNIEVRAR	540	
Db	481	LEGALKLKEISYTHAEYAAAGELKHGFLALIDADMPVIVVAPNNLEKLSKSNIEVRAR	540	
QY	541	GGQLYVFADQDAGFVSSDNNHIIEMPHVEEVIAPIFTVPLQLLAYHVALIKGTDVDPQR	600	
Db	541	GGQLYVFADQDAGFVSSDNNHIIEMPHVEEVIAPIFTVPLQLLAYHVALIKGTDVDPQR	600	
QY	601	NLAKSVTVE 609		
Db	601	NLAKSVTVE 609		
RESULT 6				
ADI38865	ADI38865 standard; protein; 609 AA.			
XX	AC	ADI38865;		
XX	AC	ADI38865;		
XX	DT	15-APR-2004 (first entry)		
XX	XX	Mutant glucosamine-6-phosphate synthase, SEQ ID 10.		
DE	XX	Glucosamine; N-acetylglucosamine; fermentation;		
KW	KW	glucosamine-6-phosphate acetyltransferase;		
KW	KW	glucosamine-6-phosphate synthase; glucosamine-6-phosphate deaminase;		
KW	KW	glucosamine-1 phosphate N-acetyltransferase; glucosamine-6-phosphate;		
KW	KW	glucosamine-1-phosphate; N-acetylglucosamine-1-phosphate;		
XX	KW	N-acetylglucosamine-6-phosphate; enzyme.		
XX	OS	Escherichia coli.		
OS	OS	Synthetic.		
XX	XX	WO2004003175-A2.		
PN	XX	08-JAN-2004.		
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			
XX	XX			

QY 541 GGQLYVFADQAGFVSSDNMHIIEMPHVEEVIAPFYTVPLQLLAYHVALIKGTDVDQPR 600  
 DB 541 GGQLYVFADQAGFVSSDNMHIIEMPHVEEVIAPFYTVPLQLLAYHVALIKGTDVDQPR 600  
 QY 601 NLAKSVTVE 609  
 DB 601 NLAKSVTVE 609  
 RESULT 7  
 ADI38867  
 ID ADI38867 standard; protein; 609 AA.  
 AC ADI38867;  
 DT 15-APR-2004 (first entry)  
 DE Mutant glucosamine-6-phosphate synthase, SEQ ID 12.  
 KW Glucosamine; N-acetylglucosamine; fermentation;  
 KW glucosamine-6-phosphate acetyltransferase;  
 KW glucosamine-6-phosphate synthase; glucosamine-6-phosphate deaminase;  
 KW glucosamine-1-phosphate N-acetyltransferase; glucosamine-6-phosphate;  
 KW glucosamine-1-phosphate; N-acetylglucosamine-1-phosphate;  
 KW N-acetylglucosamine-6-phosphate; enzyme.  
 OS Escherichia coli.  
 OS Synthetic.  
 PN WO2004/003175-A2.  
 XX 08-JAN-2004.  
 XX 01-JUL-2003; 2003WO-US020925.  
 PF 01-JUL-2002; 2002US-0393348P.  
 PR (ARKI-) ARKION LIFE SCI LLC.  
 PA Deng M, Angerer JD, Cyron D, Grund AD, Jerrell TA, Leanna C;  
 PI Mathre O, Rosson R, Running J, Severson D, Song L, Wassink S;  
 DR WPI: 2004-203380/19.  
 DR N-PSDB; ADI38866.  
 XX Producing glucosamine or N-acetylglucosamine by fermentation involves  
 PT culturing microorganism comprising glucosamine-6-phosphate  
 PT acetyltransferase, in fermentation medium, and collecting product.  
 XX Claim 15; SEQ ID NO 12; 327pp; English.  
 CC The present invention relates to a method (M1) for producing glucosamine  
 CC and N-acetylglucosamine by fermentation. The method comprises (a)  
 CC culturing in a fermentation medium a microorganism (I) which comprises  
 CC endogenous glucosamine-6-phosphate acetyltransferase (II) and a genetic  
 CC modification that increases the activity of (II), glucosamine-6-phosphate  
 CC synthase (III) or glucosamine-6-phosphate deaminase (IV), or decreases  
 CC the activity of (IV) and increases the activity of glucosamine-1  
 CC phosphate N-acetyltransferase (V), and (b) and collecting the product,  
 CC which is chosen from the group consisting of glucosamine-6-phosphate, N-  
 CC glucosamine, glucosamine-1-phosphate, N-acetylglucosamine-1-phosphate, N-  
 CC acetylglucosamine-6-phosphate, and N-acetylglucosamine. The present  
 CC sequence was used to illustrate the method of the invention.  
 XX Sequence 609 AA;  
 SQ  
 Query Match 99.4%; Score 3072; DB 8; Length 609;  
 Best Local Similarity 99.7%; Pred. No. 9.7e-270;  
 Matches 607; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 MCGIIGAIAQDVABILLLEGLRLLEYRGYDSAGLAVDTEGHMTLRLRLGKVQMLAQAAE 60  
 DB 1 MCGIIGAIAQDVABILLLEGLRLLEYRGYDSAGLAVDTEGHMTLRLRLGKVQMLAQAAE 60

QY 61 EHPHGGTGIAHTRWATHGEPSEVNAHPHVSEHIVVHNGIIEENHEPUELKARGYTFV 120  
 DB 61 EHPHGGTGIAHTRWATHGEPSEVNAHPHVSEHIVVHNGIIEENHEPUELKARGYTFV 120  
 QY 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPQLRGAYGTVIMDSRHPDPTLLAARSGSPLV 180  
 DB 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPQLRGAYGTVIMDSRHPDPTLLAARSGSPLV 180  
 QY 181 IGLGNGENFIASDQALLPVTRRFIFLEEGDIAETTRRSVNI FDKTGAEVKRDQIESNLQ 240  
 DB 181 IGLGNGENFIASDQALLPVTRRFIFLEEGDIAETTRRSVNI FDKTGAEVKRDQIESNLQ 240  
 QY 241 YDAGDKGIYHYMQKEIYEQPNAIKNTLTGRISHGQVDLSLGNADDELLSKVEHIQILA 300  
 DB 241 YDAGDKGIYHYMQKEIYEQPNAIKNTLTGRISHGQVDLSLGNADDELLSKVEHIQILA 300  
 QY 301 CGTSYNSGMVSRYPFESLAGIPCDVEIASEPRYKSAVRRNSLMITLSQSGETADTLA 360  
 DB 301 CGTSYNSGMVSRYPFESLAGIPCDVEIASEPRYKSAVRRNSLMITLSQSGETADTLA 360  
 QY 361 RLSKELGYLGSGLAICNVPSSLVRESDLALMTNAGTEIGVASTKRAFTTQLTVLLMLVAKL 420  
 DB 361 RLSKELGYLGSGLAICNVPSSLVRESDLALMTNAGTEIGVASTKRAFTTQLTVLLMLVAKL 420  
 QY 421 SRLKGLDASIEHDIVHGLQALPSRIEQMLSDQKRIEALAEAFSDKHHLFLSRGQYPIA 480  
 DB 421 SRLKGLDASIEHDIVHGLQALPSRIEQMLSDQKRIEALAEAFSDKHHLFLSRGQYPIA 480  
 QY 481 LEGALKUKELSYTHAEAYAGELKHGLALDADMPVIVVAPNNELLEKLKSNTEEVAR 540  
 DB 481 LEGALKUKELSYTHAEAYAGELKHGLALDADMPVIVVAPNNELLEKLKSNTEEVAR 540  
 QY 541 GGQLYVFADQAGFVSSDNMHIIEMPHVEEVIAPFYTVPLQLLAYHVALIKGTDVDQPR 600  
 DB 541 GGQLYVFADQAGFVSSDNMHIIEMPHVEEVIAPFYTVPLQLLAYHVALIKGTDVDQPR 600  
 QY 601 NLAKSVTVE 609  
 DB 601 NLAKSVTVE 609  
 RESULT 8  
 AAY58822  
 ID AAY58822 standard; protein; 609 AA.  
 AC AAY58822;  
 XX 08-MAY-2000 (first entry)  
 DE E. coli glucosamine-6-phosphate synthase.  
 KW Glucosamine-6-phosphate synthase; GlcN6P; glucosamine;  
 KW metabolic engineering; plasmid pKLN23-28.  
 OS Escherichia coli.  
 XX WO200004182-A1.  
 PD 27-JAN-2000.  
 PF 15-JUL-1999; 99WO-US015976.  
 XX 15-JUL-1998; 98US-00115475.  
 PA (DCVB-) DCV INC DBA BIO-TECH RESOURCES.  
 XX Berry A, Burlingame RP, Millis JR;  
 DR WPI; 2000-182441/16.  
 DR N-PSDB; AAZ58249.  
 PT Fermentation of E. coli having an altered amino acid sugar metabolic

pathway to produce glucosamine, especially using novel recombinant variant glucosamine-6-phosphate synthases.

Claim 15; Page 111-113; 150pp; English.

The present sequence is that of wild-type glucosamine-6-phosphate synthase (GlcN6P synthase) of *Escherichia coli* strain W3110. The sequence is predicted from the isolated glms gene (see AA258249) in plasmid pGKN23 -28. Recombinant nucleic acids encoding GlcN6P synthase are claimed, and are used for the expression of the enzyme in host microbial strains, especially *E. coli*, for use in the production of glucosamine. The invention also provides methods for the overproduction of glucosamine using a genetically engineered microorganism that encodes a GlcN6P synthase modified to increase its activity (see also AA58823-27). Preferred modifications comprise amino acid deletion, insertion, inversion, derivatisation or substitution, especially I4T, I272T, S240P, A39T, R250C, G472S and L469P substitution, of the present sequence

XX Sequence 609 AA;

Query Match 99.3%; Score 3068; DB 3; Length 609;  
 Best Local Similarity 99.5%; Pred. No. 2.2e-269;  
 Matches 606; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MCGIVGAIQRDVAEILLEGRLRLEYRGYDSAGLAVVDTEGHMTRRLRGKVMQLAAAE 60  
 Db 1 MCGIVGAIQRDVAEILLEGRLRLEYRGYDSAGLAVVDTEGHMTRRLRGKVMQLAAAE 60

Qy 61 EPHLHGSGTGAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHPELREELKARGYTFV 120  
 Db 61 EPHLHGSGTGAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHPELREELKARGYTFV 120

Qy 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPOLRGAYGTVMDSRHPDTLAARGSPPLV 180  
 Db 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPOLRGAYGTVMDSRHPDTLAARGSPPLV 180

Qy 181 IGLGMGNFTASDQALLPVTFRFIFLEBGDIAETRRSVNIFDKTGAEVKRDIESNLQ 240  
 Db 181 IGLGMGNFTASDQALLPVTFRFIFLEBGDIAETRRSVNIFDKTGAEVKRDIESNLQ 240

Qy 241 YDAGDGIYCHYMQKEIYQPNAIKNTLTGRISHGQVDISELGPNADELLSKVEHIQILA 300  
 Db 241 YDAGDGIYCHYMQKEIYQPNAIKNTLTGRISHGQVDISELGPNADELLSKVEHIQILA 300

Qy 301 CGFSYNGMSVRWYFESLAGIPCDVEIASEFRRYSKSAVRNSIMITLSQSGTADTLA 360  
 Db 301 CGFSYNGMSVRWYFESLAGIPCDVEIASEFRRYSKSAVRNSIMITLSQSGTADTLA 360

Qy 361 RLSKELGYLGLSLAICNVPQSSLVRESDLMTNAGTEIGVASTKFTQTLTVLLMLVAKL 420  
 Db 361 RLSKELGYLGLSLAICNVPQSSLVRESDLMTNAGTEIGVASTKFTQTLTVLLMLVAKL 420

Qy 421 SRLKGLDASIEHDIYHGLQALPSRIQMLISQKRIEALAEFSDKHHLFLSRGQOYPIA 480  
 Db 421 SRLKGLDASIEHDIYHGLQALPSRIQMLISQKRIEALAEFSDKHHLFLSRGQOYPIA 480

Qy 481 LEGALKLKEISYTHAEYAAAGELKHGFLALIDADMPVIVVAPNNELLEKLKSNIEVVR 540  
 Db 481 LEGALKLKEISYTHAEYAAAGELKHGFLALIDADMPVIVVAPNNELLEKLKSNIEVVR 540

Qy 541 GGQLYVFADQDAGFVSDNNHIITEMPHVEVIAPIFYTVPLQLLAYHVALIKGTVDQPR 600  
 Db 541 GGQLYVFADQDAGFVSDNNHIITEMPHVEVIAPIFYTVPLQLLAYHVALIKGTVDQPR 600

Qy 601 NLAKSVTVE 609  
 Db 601 NLAKSVTVE 609

RESULT 9  
 AAU34806  
 ID AAU34806 standard; protein; 609 AA.  
 XX

Db 121 SETDTEVIAHLVNWELKOGGTLREAVLRAIPQLRGAYGTVMDSRHPDTLLAARSGSPLV 180  
Qy 181 IGLGMENFIASDQALLPVTRRFIFLEEGDIAETIRRSVNI FDKTGAEVKRDIESNLQ 240  
Db 181 IGLGMENFIASDQALLPVTRRFIFLEEGDIAETIRRSVNI FDKTGAEVKRDIESNLQ 240  
Qy 241 YDAGDKGIYHYMOKEIYEOPNAIKNTLTGRISHGQVDLSLGNADLLSKVEHIQILA 300  
Db 241 YDAGDKGIYHYMOKEIYEOPNAIKNTLTGRISHGQVDLSLGNADLLSKVEHIQILA 300  
Qy 301 CGTSYNSGMVSRYPFESLAGIPCDVEIASFPRYKSAVRNSLMTLSQSGETADTLA 360  
Db 301 CGTSYNSGMVSRYPFESLAGIPCDVEIASFPRYKSAVRNSLMTLSQSGETADTLA 360  
Qy 361 RLSKELGYLGLSLAICNVPGSSLVRESLALMTNAGTEIGVASTKFTTQTLVLLMLVAKL 420  
Db 361 RLSKELGYLGLSLAICNVPGSSLVRESLALMTNAGTEIGVASTKFTTQTLVLLMLVAKL 420  
Qy 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEFSDKHAFILSRGDOYPIA 480  
Db 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEFSDKHAFILSRGDOYPIA 480  
Qy 481 LEGALKLKEISYIHAEVAAGELKHGPLALIDADMPVIWVAPNNLEKLSNIEVRAR 540  
Db 481 LEGALKLKEISYIHAEVAAGELKHGPLALIDADMPVIWVAPNNLEKLSNIEVRAR 540  
Qy 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIPIFYTVPLQLLAYHVALIKGTVDQDPR 600  
Db 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIPIFYTVPLQLLAYHVALIKGTVDQDPR 600  
Qy 601 NLAKSVTVE 609  
Db 601 NLAKSVTVE 609

RESULT 10  
ABU28829  
ID ABU28829 standard; protein; 609 AA.  
XX  
AC ABU28829;  
DT 19-JUN-2003 (first entry)  
XX  
DE Protein encoded by Prokaryotic essential gene #14356.  
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.  
XX  
OS Escherichia coli.  
XX  
PN WO20027183-A2.  
XX  
PD 03-OCT-2002.  
XX  
PF 21-MAR-2002; 2002WO-US009107.  
XX  
PR 21-MAR-2001; 2001US-00815242.  
PR 06-SEP-2001; 2001US-00948993.  
PR 25-OCT-2001; 2001US-0342923P.  
PR 08-FEB-2002; 2002US-00072851.  
PR 06-MAR-2002; 2002US-0362699P.  
XX  
PA (ELIT-) ELITRA PHARM INC.  
XX  
PI Wang L, Zamudio C, Malone C, Haseibeck R, Ohlsen KL, Zyskind JW;  
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
XX  
DR WPI: 2003-029926/02.  
DR N-PSDB; AC32699.  
XX  
XX New antisense nucleic acids, useful for identifying proteins or screening  
PT for homologous nucleic acids required for cellular proliferation to  
PT isolate candidate molecules for rational drug discovery programs.  
XX

PS Claim 25; SEQ ID NO 56753; 1766pp; English.  
XX  
CC The invention relates to an isolated nucleic acid comprising any one of  
CC the 6213 antisense sequences given in the specification where expression  
CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
CC (1) a vector comprising a promoter operably linked to the nucleic acid  
CC encoding a polypeptide whose expression is inhibited by the antisense  
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
CC polypeptide or its fragment whose expression is inhibited by the  
CC antisense nucleic acid; (4) an antibody capable of specifically binding  
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
CC proliferation or the activity of a gene in an operon required for  
CC proliferation; (7) identifying a compound that influences the activity of  
CC the gene product or that has an activity against a biological pathway  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC or a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
CC the target prokaryotic essential genes. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences  
SQ Sequence 609 AA;

Query Match 99.3%; Score 3068; DB 6; Length 609;  
Best Local Similarity 99.5%; Pred. No. 2.2e-269;  
Matches 606; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
Qy 1 MCGIVGAIAQRDVAEILLEGRLRLEYRGYDSAGLAVVDTEGHMTRLRRLGKVMQAAAE 60  
Db 1 MCGIVGAIAQRDVAEILLEGRLRLEYRGYDSAGLAVVDTEGHMTRLRRLGKVMQAAAE 60  
Qy 61 EHPHGGTGIAHTWATHGEPSEVNAHPHVEHIVVHNGIIEHPELREELKARGYTFV 120  
Db 61 EHPHGGTGIAHTWATHGEPSEVNAHPHVEHIVVHNGIIEHPELREELKARGYTFV 120  
Qy 121 SETDTEVIAHLVNWELKOGGTLREAVLRAIPQLRGAYGTVMDSRHPDTLLAARSGSPLV 180  
Db 121 SETDTEVIAHLVNWELKOGGTLREAVLRAIPQLRGAYGTVMDSRHPDTLLAARSGSPLV 180  
Qy 181 IGLGMENFIASDQALLPVTRRFIFLEEGDIAETIRRSVNI FDKTGAEVKRDIESNLQ 240  
Db 181 IGLGMENFIASDQALLPVTRRFIFLEEGDIAETIRRSVNI FDKTGAEVKRDIESNLQ 240  
Qy 241 YDAGDKGIYHYMOKEIYEOPNAIKNTLTGRISHGQVDLSLGNADLLSKVEHIQILA 300  
Db 241 YDAGDKGIYHYMOKEIYEOPNAIKNTLTGRISHGQVDLSLGNADLLSKVEHIQILA 300  
Qy 301 CGTSYNSGMVSRYPFESLAGIPCDVEIASFPRYKSAVRNSLMTLSQSGETADTLA 360  
Db 301 CGTSYNSGMVSRYPFESLAGIPCDVEIASFPRYKSAVRNSLMTLSQSGETADTLA 360  
Qy 361 RLSKELGYLGLSLAICNVPGSSLVRESLALMTNAGTEIGVASTKFTTQTLVLLMLVAKL 420  
Db 361 RLSKELGYLGLSLAICNVPGSSLVRESLALMTNAGTEIGVASTKFTTQTLVLLMLVAKL 420  
Qy 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEFSDKHAFILSRGDOYPIA 480  
Db 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEFSDKHAFILSRGDOYPIA 480  
Qy 481 LEGALKLKEISYIHAEVAAGELKHGPLALIDADMPVIWVAPNNLEKLSNIEVRAR 540

Db 481 LEGALKKEISYHAEYAAAGELKHGFLALIDAMPVIVVAPNNELEKUKSIEVRAR 540  
 QY 541 GGQLYVFADQDAGFVSSDNHIIEMPHVEEVIPIFYTVPLQLLAYHVALIKGTVDQPR 600  
 Db 541 GGQLYVFADQDAGFVSSDNHIIEMPHVEEVIPIFYTVPLQLLAYHVALIKGTVDQPR 600  
 QY 601 NLAQSVTVE 609  
 Db 601 NLAQSVTVE 609  
 RESULT 11  
 ADI38857 standard; protein; 609 AA.  
 AC ADI38857;  
 DT 15-APR-2004 (first entry)  
 DE Glucosamine-6-phosphate synthase, glms, SEQ ID 2.  
 KW Glucosamine; N-acetylglucosamine; fermentation;  
 KW glucosamine-6-phosphate acetyltransferase;  
 KW glucosamine-6-phosphate synthase; glucosamine-6-phosphate deaminase;  
 KW glucosamine-1 phosphate N-acetyltransferase; glucosamine-6-phosphate;  
 KW glucosamine-1-phosphate; N-acetylglucosamine-1-phosphate;  
 KW N-acetylglucosamine-6-phosphate; enzyme.  
 OS Escherichia coli.  
 XX WO2004003175-A2.  
 XX 08-JAN-2004.  
 PF 01-JUL-2003; 2003WO-US020925.  
 PR 01-JUL-2002; 2002US-0393348P.  
 XX (ARKI-) ARKION LIFE SCI LLC.  
 PI Deng M, Angerer JD, Cyton D, Grund AD, Jerrell TA, Leanna C;  
 PI Mathre O, Rosson R, Running J, Severson D, Song L, Wassink S;  
 XX WPI; 2004-203380/19.  
 DR N-PSDB; ADI38856.  
 XX Producing glucosamine or N-acetylglucosamine by fermentation involves  
 PT culturing microorganism comprising glucosamine-6-phosphate  
 PT acetyltransferase, in fermentation medium, and collecting product.  
 XX Claim 15; SEQ ID NO 2; 327pp; English.  
 CC The present invention relates to a method (M1) for producing glucosamine  
 CC and N-acetylglucosamine by fermentation. The method comprises (a)  
 CC culturing in a fermentation medium a microorganism (I) which comprises  
 CC endogenous glucosamine-6-phosphate acetyltransferase (II) and a genetic  
 CC modification that increases the activity of (III), glucosamine-6-phosphate  
 CC synthase (III) or glucosamine-6-phosphate deaminase (IV), or decreases  
 CC the activity of (IV) and increases the activity of glucosamine-1  
 CC phosphate N-acetyltransferase (V), and (b) and collecting the product,  
 CC which is chosen from the group consisting of glucosamine-6-phosphate, N-  
 CC glucosamine, glucosamine-1-phosphate, N-acetylglucosamine-1-phosphate, N-  
 CC acetylglucosamine-6-phosphate, and N-acetylglucosamine. The present  
 CC sequence was used to illustrate the method of the invention.  
 XX Sequence 609 AA;  
 Query Match 99.3%; Score 3068; DB 8; Length 609;  
 Best Local Similarity 99.5%; Pred. No. 2.2e-269;  
 Matches 606; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 MCGIVGAIQARDVAEILLEGRLRYRGYDSAGLAVVDTEGHMTRRLRKGVQMLQAQAE 60

Db 1 MCGIVGAIQARDVAEILLEGRLRYRGYDSAGLAVVDTEGHMTRRLRKGVQMLQAQAE 60  
 QY 61 EHPHGGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHNPPLREELKARGYTFV 120  
 Db 61 EHPHGGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHNPPLREELKARGYTFV 120  
 QY 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPOLRGAYGTVMDSRHPDITLAARSQPLV 180  
 Db 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPOLRGAYGTVMDSRHPDITLAARSQPLV 180  
 QY 181 IGLMGENTIASDQALLPVTRRFIFLEBGDIAEITRRSVNI FDKTCAEVKRODIESNLQ 240  
 Db 181 IGLMGENTIASDQALLPVTRRFIFLEBGDIAEITRRSVNI FDKTCAEVKRODIESNLQ 240  
 QY 241 YDAGDKGIYRHYMQKEIYEOPNAIKNTLTGRISHGQVDLSELGNADDELLSKVEHIQILA 300  
 Db 241 YDAGDKGIYRHYMQKEIYEOPNAIKNTLTGRISHGQVDLSELGNADDELLSKVEHIQILA 300  
 QY 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASSEFRYKSAVRNSLIMITLSQSGETADTLA 360  
 Db 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASSEFRYKSAVRNSLIMITLSQSGETADTLA 360  
 QY 361 RLKSELGYLGLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQTLVLLMLVAKL 420  
 Db 361 RLKSELGYLGLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQTLVLLMLVAKL 420  
 QY 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEFSDKHHLALFLSRGQDQPIA 480  
 Db 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEFSDKHHLALFLSRGQDQPIA 480  
 QY 481 LEGALKKEISYHAEYAAAGELKHGFLALIDAMPVIVVAPNNELEKUKSIEVRAR 540  
 Db 481 LEGALKKEISYHAEYAAAGELKHGFLALIDAMPVIVVAPNNELEKUKSIEVRAR 540  
 QY 541 GGQLYVFADQDAGFVSSDNHIIEMPHVEEVIPIFYTVPLQLLAYHVALIKGTVDQPR 600  
 Db 541 GGQLYVFADQDAGFVSSDNHIIEMPHVEEVIPIFYTVPLQLLAYHVALIKGTVDQPR 600  
 QY 601 NLAQSVTVE 609  
 Db 601 NLAQSVTVE 609  
 RESULT 12  
 ADS45181 standard; protein; 609 AA.  
 ID ADS45181  
 AC ADS45181;  
 DT 02-DEC-2004 (first entry)  
 DE Bacterial polypeptide #23611.  
 KW Recombinant DNA construct; transformed plant; improved plant property;  
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;  
 KW pathogen tolerance; pest tolerance; plant disease resistance;  
 KW cell cycle pathway modification; plant growth regulator;  
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;  
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
 KW bacterial polypeptide.  
 XX Bacteria.  
 OS US2003233675-A1.  
 XX 18-DEC-2003.  
 PD 20-FEB-2003; 2003US-00369493.  
 PF 21-FEB-2002; 2002US-0360039P.  
 XX (CAOY/) CAO Y.

PA (HINK/) HINKLE G J.  
PA (SLAT/) SLATER S C.  
PA (CHEN/) CHEN X.  
PA (GOLD/) GOLDMAN B S.  
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;  
XX WPI; 2004-061375/06.  
DR New recombinant DNA construct comprising a promoter positioned to provide  
PT for expression of a polynucleotide encoding a polypeptide from a  
PT microbial source, useful for producing plants with improved properties.  
XX  
PS Claim 1; SEQ ID NO 23611; 122pp; English.  
XX  
CC The invention relates to a recombinant DNA construct comprising a  
CC promoter functional in a plant cell, where the promoter is positioned to  
CC provide for expression of a polynucleotide encoding a polypeptide from a  
CC microbial source. The invention also relates to a transformed plant  
CC comprising the recombinant DNA construct and a method of producing a  
CC transformed plant having an improved property. The plant is a crop plant  
CC such as maize or soybean. The method of producing a transformed plant  
CC having an improved property comprises transforming a plant with the  
CC recombinant DNA construct and growing the transformed plant, where the  
CC polynucleotide or polypeptide is useful for improving plant properties.  
CC The recombinant DNA construct is useful for producing plants with  
CC improved plant properties, e.g. improved cold, heat or drought tolerance,  
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
CC increased resistance to plant disease, better growth rate by modification  
CC of the cell cycle pathway with plant growth regulators, increased rate of  
CC homologous recombination, modified seed oil or protein yield and/or  
CC content, improved yield by modification of carbohydrate, nitrogen or  
CC phosphorus use and/or uptake, by modification of photosynthesis or by  
CC providing improved plant growth and development under at least one stress  
CC condition, improved lignin production or improved galactomannan  
CC production. This sequence represents a bacterial polypeptide used in the  
CC scope of the invention. Note: The sequence data for this patent did not  
CC form part of the printed specification but was obtained in electronic  
XX format from USPTO at seqdata.uspto.gov/sequence.html.  
SQ Sequence 609 AA;  
Query Match 99.3%; Score 3068; DB 8; Length 609;  
Best Local Similarity 99.5%; Pred. No. 2.2e-263;  
Matches 606; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 MCGIVGAIQARDVAILLEGRLRLEYRGYDAGLAVDTEGHMTRRLRLGKVMQAQAAE 60  
DB 1 MCGIVGAIQARDVAILLEGRLRLEYRGYDAGLAVDTEGHMTRRLRLGKVMQAQAAE 60  
QY 61 EPHLGGTGIATRWATHGEPSEVNAHPVSEHIVVHNGIENHPELREELKARGYTFV 120  
DB 61 EPHLGGTGIATRWATHGEPSEVNAHPVSEHIVVHNGIENHPELREELKARGYTFV 120  
QY 121 SETDTEVIAHLVNNELKOGGTLREAVLRAIPQLRGAYGTVMDSRHPDTLLAARSGSLV 180  
DB 121 SETDTEVIAHLVNNELKOGGTLREAVLRAIPQLRGAYGTVMDSRHPDTLLAARSGSLV 180  
QY 181 IGLGMGENFIASDQALALPVTRRRPFLREGDIAETITRSVNI FDKTGAEVKRDITSLQ 240  
DB 181 IGLGMGENFIASDQALALPVTRRRPFLREGDIAETITRSVNI FDKTGAEVKRDITSLQ 240  
QY 241 YDAGDKGIYCHMYQKEIYEQNAIKNTLTGRISHGQVDLSBELGNADLLSKVEHIQILA 300  
DB 241 YDAGDKGIYCHMYQKEIYEQNAIKNTLTGRISHGQVDLSBELGNADLLSKVEHIQILA 300  
QY 301 CGTNSNMGVSRYPFESLAGIPCDVEIASERPRKSAVERNSLMTITLSQGETADTLAGL 360  
DB 301 CGTNSNMGVSRYPFESLAGIPCDVEIASERPRKSAVERNSLMTITLSQGETADTLAGL 360  
QY 361 RLSKELGYLGLSLAICNVPGSSILVRESDLALMTNAGTEIGVASTKRAFTTQLTVLLMLVAKL 420  
DB 361 RLSKELGYLGLSLAICNVPGSSILVRESDLALMTNAGTEIGVASTKRAFTTQLTVLLMLVAKL 420

QY 421 SRLKGLDASIEHDIHVHGLQALPSRIEQMLSDQKRIEALAEFSDKHAFILSRGQDYPIA 480  
DB 421 SRLKGLDASIEHDIHVHGLQALPSRIEQMLSDQKRIEALAEFSDKHAFILSRGQDYPIA 480  
QY 481 LEGALKUKEISYIHAEAYAGELKHGPLALDADMPVIVVAPNNELLEKLKSNIEEVRAR 540  
DB 481 LEGALKUKEISYIHAEAYAGELKHGPLALDADMPVIVVAPNNELLEKLKSNIEEVRAR 540  
QY 541 GGQLYVFADQDAGFVSSDNMHIIEMPHVEEVIAPFYTPVQLLAYHVALIKGTDVDDQPR 600  
DB 541 GGQLYVFADQDAGFVSSDNMHIIEMPHVEEVIAPFYTPVQLLAYHVALIKGTDVDDQPR 600  
QY 601 NLAKSVTVE 609  
DB 601 NLAKSVTVE 609  
RESULT 13  
ADU00590  
ID ADU00590 standard; protein; 609 AA.  
XX AC ADU00590;  
XX 13-JAN-2005 (first entry)  
DT Amino acid sequence of a glucosamine-6-phosphate synthase.  
DE chitin; chitosan; fermentation;  
XX glutamine-fructose-6-phosphate amidotransferase; glucosamine synthetase;  
KW glucosamine-6-phosphate synthase; GNA1 gene; glms gene;  
KW glucosamine-6-phosphate acetyltransferase; GNA1 gene; chitin synthase;  
KW chitin deacetylase; CDA1 gene; CDA2 gene;  
KW N-acetylglucosamine-6-phosphate deacetylase;  
KW glucosamine-6-phosphate deaminase; chitinase; chitosanase; fungus; yeast;  
XX enzyme.  
OS Escherichia coli.  
XX WO2004092391-A2.  
XX 28-OCT-2004.  
XX 12-APR-2004; 2004WO-US011286.  
XX 11-APR-2003; 2003US-0462087P.  
XX (ARKI-) ARKION LIFE SCI LLC.  
FI Deng M, McMullin TW, Grund AD;  
XX WPI; 2004-766880/75.  
DR N-PSDB; ADU00589.  
XX Producing chitin or chitosan, involves culturing microorganism with  
PT genetic modification that results in increase in activity of glutamine-  
PT fructos-6-phosphate amidotransferase, in fermentation medium, to produce  
PT chitin or chitosan.  
XX Example 1; SEQ ID NO 23; 161pp; English.  
XX The specification describes a method for producing chitin or chitosan by  
CC a fermentation process. The method involves culturing in a fermentation  
CC medium a microorganism which comprises one or more genetic modifications  
CC that result in an increase in the activity of glutamine-fructose-6-  
CC phosphate amidotransferase (also known as glucosamine synthetase and  
CC glucosamine-6-phosphate synthase, and encoded by the GNA1 eukaryotic gene  
CC and the glms bacterial gene), glucosamine-6-phosphate acetyltransferase  
CC (encoded by GNA1), chitin synthase or chitin deacetylase (encoded by CDA1  
CC and CDA2), or in decrease in the activity of N-acetylglucosamine-6-  
CC phosphate deacetylase, glucosamine-6-phosphate deaminase, chitinase and  
CC chitosanase, and collecting the chitin or chitosan. The method is useful  
CC for producing chitin or chitosan by utilizing microorganisms such as



CC fungus, yeast (e.g. Saccharomyces or Schizosaccharomyces ) and  
CC filamentous fungus (e.g. Aspergillus, Absidia or Rhizopus), preferably S.  
CC cerevisiae, A. niger or A. nidulans. The method enables high quantities  
CC of chitin and chitosan to be produced cost effectively. The present  
CC sequence represents an Escherichia coli glucosamine-6-phosphate synthase.  
CC It was used to transform yeast for use in the method of the invention.  
XX  
XX  
SQ Sequence 609 AA;

Query Match		99.3%;	Score 3068;	DB 8;	Length 609;
Best Local Similarity		99.5%;	Pred. No. 2.2e-269;		
Matches 606;		Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;
QY	1	MCGIVGAIQRDVAEILLGLRLRYGVDSAGLAVVDTGHWRLRLRLKGVQMLAAAE	60		
Db	1	MCGIVGAIQRDVAEILLGLRLRYGVDSAGLAVVDTGHWRLRLRLKGVQMLAAAE	60		
QY	61	EHLPGGTGIAHTRWATHGEPSEVNAHPVSEHVVVHNGIIEHPEELKARGYTFV	120		
Db	61	EHLPGGTGIAHTRWATHGEPSEVNAHPVSEHVVVHNGIIEHPEELKARGYTFV	120		
QY	121	SETDTEVIAHLVNWELKQGGTLREAVLRAIPOLRGAYGTVMDSRHPDITLLAARSGPLV	180		
Db	121	SETDTEVIAHLVNWELKQGGTLREAVLRAIPOLRGAYGTVMDSRHPDITLLAARSGPLV	180		
QY	181	IGLWGENFIASDQALLPVTRRFIFLEGGDIAETIRRSVNIPTKGAEVKQDIESNLQ	240		
Db	181	IGLWGENFIASDQALLPVTRRFIFLEGGDIAETIRRSVNIPTKGAEVKQDIESNLQ	240		
QY	241	YDAGKGIYCHYMOKIYEOPNAIKNTLTGRISHGQVDLSGLPNADELLSKVEHIQILA	300		
Db	241	YDAGKGIYCHYMOKIYEOPNAIKNTLTGRISHGQVDLSGLPNADELLSKVEHIQILA	300		
QY	301	CCTSNGMVSRYWFESLAGIPCDVEIASEFRYKSAVRNSLMTLSQGETADTLAGL	360		
Db	301	CCTSNGMVSRYWFESLAGIPCDVEIASEFRYKSAVRNSLMTLSQGETADTLAGL	360		
QY	361	LSKELGYSGLAIACNVPSSLVRESDLALMTNAGTEIGVASTKFTTQTLVLLMLVAKL	420		
Db	361	LSKELGYSGLAIACNVPSSLVRESDLALMTNAGTEIGVASTKFTTQTLVLLMLVAKL	420		
QY	421	SRLKGLDASIEHDIHVHGLQALPSRIEQLMSQDKRIEALAEDEFSDKHAFFLSRGDOYPIA	480		
Db	421	SRLKGLDASIEHDIHVHGLQALPSRIEQLMSQDKRIEALAEDEFSDKHAFFLSRGDOYPIA	480		
QY	481	LEGALKLKEISYTHAEYAAAGELKHGFLALIDADMPVIVVAPNNELLEKKNIEVRAR	540		
Db	481	LEGALKLKEISYTHAEYAAAGELKHGFLALIDADMPVIVVAPNNELLEKKNIEVRAR	540		
QY	541	GGQLYVFADODAGFVSSDNHIIEMPHVEVIAPIETVPLQLLAYHVALIKGTVDQPR	600		
Db	541	GGQLYVFADODAGFVSSDNHIIEMPHVEVIAPIETVPLQLLAYHVALIKGTVDQPR	600		
QY	601	NLAKSVTVE 609			
Db	601	NLAKSVTVE 609			

RESULT 14  
ADW23840  
ID ADW23840 standard; protein; 608 AA.  
XX  
XX  
AC ADW23840;  
XX  
XX  
DT 07-APR-2005 (first entry)  
XX  
XX  
DE Novel human GFAT protein-related E coli GFAT protein SeqID13.  
XX  
XX  
KW protein purification; glutamine:fructose-6-phosphate amidotransferase;  
KW GFAT; antidiabetic; anorectic; metabolic; cytosolic; osteopathic;  
KW fungicide; herbicide; diabetes; non-insulin dependent diabetes; obesity;  
KW acidosis; cancer; osteoporosis.

OS Escherichia coli.  
XX  
XX  
PN FR2857374-Al.  
XX  
XX  
PD 14-JAN-2005.  
XX  
XX  
PF 08-JUL-2003; 2003PR-00008350.  
XX  
XX  
PR 08-JUL-2003; 2003PR-00008350.  
XX  
XX  
PA (CNRS ) CNRS CENT NAT RECH SCI.  
XX  
XX  
PI Badet DMAJE, Badet BF;  
XX  
XX  
XX WPI; 2005-076419/09.  
XX  
XX  
PT New modified glutamine:fructose-6-phosphate amidotransferase, useful for  
PT identifying specific inhibitors, potentially useful for treating e.g.  
PT diabetes, contains a purification tag, also related nucleic acid.  
XX  
XX  
PS Disclosure; SEQ ID NO 13; 61pp; French.  
XX  
XX  
CC This invention relates to a novel protein corresponding to human  
CC enzymatically active but modified glutamine:fructose-6-phosphate  
CC amidotransferase (GFAT). The invention may be useful for the production  
CC of compounds with an antidiabetic, anorectic, metabolic, cytosolic,  
CC osteopathic, fungicide or herbicide activity. The invention may be used  
CC to screen for agents that modify, especially inhibit, its activity,  
CC (particularly type II), obesity, acidosis, ketosis, cancer and  
CC osteoporosis. In addition inhibitors of plant and fungal GFAT may be  
CC useful as herbicides and fungicides, respectively. The present sequence  
CC is that of a human GFAT-like protein of the invention.

SQ Sequence 608 AA;

Query Match		99.2%;	Score 3063;	DB 9;	Length 608;
Best Local Similarity		99.5%;	Pred. No. 6.3e-269;		
Matches 605;		Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;
QY	2	CGIVGAIQRDVAEILLGLRLRYGVDSAGLAVVDTGHWRLRLRLKGVQMLAAAE	61		
Db	1	CGIVGAIQRDVAEILLGLRLRYGVDSAGLAVVDTGHWRLRLRLKGVQMLAAAE	60		
QY	62	HPLHGGTGIATRWATHGEPSEVNAHPVSEHVVVHNGIIEHPEELKARGYTFVS	121		
Db	61	HPLHGGTGIATRWATHGEPSEVNAHPVSEHVVVHNGIIEHPEELKARGYTFVS	120		
QY	122	ETDTEVIAHLVNWELKQGGTLREAVLRAIPOLRGAYGTVMDSRHPDITLLAARSGPLVI	181		
Db	121	ETDTEVIAHLVNWELKQGGTLREAVLRAIPOLRGAYGTVMDSRHPDITLLAARSGPLVI	180		
QY	182	GLGNGENFIASDQALLPVTRRFIFLEGGDIAETIRRSVNIPTKGAEVKQDIESNLQY	241		
Db	181	GLGNGENFIASDQALLPVTRRFIFLEGGDIAETIRRSVNIPTKGAEVKQDIESNLQY	240		
QY	242	DAGDKGIYCHYMOKIYEOPNAIKNTLTGRISHGQVDLSGLPNADELLSKVEHIQILAC	301		
Db	241	DAGDKGIYCHYMOKIYEOPNAIKNTLTGRISHGQVDLSGLPNADELLSKVEHIQILAC	300		
QY	302	GTSYNGMVSRYWFESLAGIPCDVEIASEFRYKSAVRNSLMTLSQGETADTLAGLR	361		
Db	301	GTSYNGMVSRYWFESLAGIPCDVEIASEFRYKSAVRNSLMTLSQGETADTLAGLR	360		
QY	362	LSKELGYSGLAIACNVPSSLVRESDLALMTNAGTEIGVASTKFTTQTLVLLMLVAKLS	421		
Db	361	LSKELGYSGLAIACNVPSSLVRESDLALMTNAGTEIGVASTKFTTQTLVLLMLVAKLS	420		
QY	422	RUKGLDASIEHDIHVHGLQALPSRIEQLMSQDKRIEALAEDEFSDKHAFFLSRGDOYPIAL	481		
Db	421	RUKGLDASIEHDIHVHGLQALPSRIEQLMSQDKRIEALAEDEFSDKHAFFLSRGDOYPIAL	480		
QY	482	EGALKLKEISYTHAEYAAAGELKHGFLALIDADMPVIVVAPNNELLEKKNIEVRARG	541		



GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 14, 2006, 15:28:50 ; Search time 209.344 Seconds  
(without alignments)  
1347.534 Million cell updates/sec

Title: US-10-612-779-6

Perfect score: 3089

Sequence: 1 MCGIVGATAQDVAEILLEG.....LIKGTVDQPRNAKSTVTE 609

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA Main:\*

- 1: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 3: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 4: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 5: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 6: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US11\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3089	100.0	609	4	US-10-024-460-22
2	3089	100.0	609	4	US-10-612-779-6
3	3089	100.0	609	5	US-10-823-397-25
4	3072	99.4	609	4	US-10-024-460-28
5	3072	99.4	609	4	US-10-024-460-31
6	3072	99.4	609	4	US-10-612-779-10
7	3072	99.4	609	4	US-10-612-779-12
8	3068	99.3	609	3	US-09-815-242-10399
9	3068	99.3	609	4	US-10-024-460-16
10	3068	99.3	609	4	US-10-369-493-23611
11	3068	99.3	609	4	US-10-282-122A-56753
12	3068	99.3	609	4	US-10-612-779-2
13	3068	99.3	609	5	US-10-823-397-23
14	3061	99.1	609	4	US-10-024-460-25
15	3061	99.1	609	4	US-10-612-779-8
16	3053	98.8	609	4	US-10-024-460-19
17	3053	98.8	609	4	US-10-612-779-4
18	3043	98.5	609	4	US-10-612-779-14
19	3035	98.3	609	3	US-09-815-242-14065
20	3035	98.3	609	4	US-10-282-122A-76152
21	2932	94.9	609	4	US-10-282-122A-56279
22	2862	92.7	609	3	US-09-815-242-11758
23	2857	92.5	609	4	US-10-282-122A-59335
24	2700.5	87.4	608	4	US-10-282-122A-72965
25	2668	86.4	609	4	US-10-282-122A-78533
26	2647	85.7	609	4	US-10-369-493-21242
27	2605	84.3	609	4	US-10-369-493-438

28	2493.5	80.7	608	4	US-10-282-122A-63028	Sequence 69028, A
29	2256.5	73.0	610	3	US-09-815-242-11018	Sequence 11018, A
30	2256.5	73.0	610	4	US-10-282-122A-S8163	Sequence 58163, A
31	2256.5	73.0	610	6	US-11-194-246-313	Sequence 313, App
32	2250.5	70.8	610	4	US-10-282-122A-77063	Sequence 77063, A
33	2188.5	70.8	610	4	US-10-282-122A-67388	Sequence 67388, A
34	1993	64.5	611	4	US-10-282-122A-67654	Sequence 67654, A
35	1990.5	64.4	610	4	US-10-369-493-13847	Sequence 13847, A
36	1983	64.2	611	3	US-09-815-242-12111	Sequence 12111, A
37	1983	64.2	611	4	US-10-282-122A-66848	Sequence 66848, A
38	1975	63.9	611	4	US-10-282-122A-69638	Sequence 69638, A
39	1797.5	58.2	612	4	US-10-282-122A-44943	Sequence 44943, A
40	1734	56.1	609	4	US-10-369-493-15424	Sequence 15424, A
41	1734	56.1	609	4	US-10-369-493-15792	Sequence 15792, A
42	1712	55.4	609	4	US-10-369-493-16174	Sequence 16174, A
43	1710.5	55.4	604	4	US-10-282-122A-61472	Sequence 61472, A
44	1691	54.7	615	4	US-10-369-493-19598	Sequence 19598, A
45	1690	54.7	609	4	US-10-369-493-9210	Sequence 9210, Ap

ALIGNMENTS

RESULT 1  
US-10-024-460-22  
; Sequence 22, Application US/10024460  
; Publication No. US20030044939A1  
; GENERAL INFORMATION:  
; APPLICANT: Berry, Alan  
; APPLICANT: Burlingame, Richard P.  
; APPLICANT: Millis, James R.  
; TITLE OF INVENTION: PROCESS AND MATERIALS FOR PRODUCTION OF GLUCOSAMINE  
; FILE REFERENCE: 3161-18-C1  
; CURRENT APPLICATION NUMBER: US/10/024,460  
; PRIOR FILING DATE: 2001-12-17  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/115,475  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-15  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/035,494  
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-01-14  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 22  
; LENGTH: 609  
; TYPE: PRT  
; ORGANISM: Escherichia coli  
; US-10-024-460-22

Query Match	100.0%	Score 3089;	DB 4;	Length 609;
Best Local Similarity	100.0%	Pred. No. 4.2e-252;		
Matches 609;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MCGIVGATAQDVAEILLEGRLRLEVRGYDSAGLAVVDTEGHMTRRLRLKQVOMLAQAAE	60	
Db	1	MCGIVGATAQDVAEILLEGRLRLEVRGYDSAGLAVVDTEGHMTRRLRLKQVOMLAQAAE	60	
QY	61	EHPHGGTGIAATRWATHGEPSEVNAHPVSHIIVVHNGIITENHEPRLKARGYTFV	120	
Db	61	EHPHGGTGIAATRWATHGEPSEVNAHPVSHIIVVHNGIITENHEPRLKARGYTFV	120	
QY	121	SETDTEVIAHLVNNWELKQGGTLREAVLRAIPOLRGAYGTVMDSRHPDTLLAARSGSPLV	180	
Db	121	SETDTEVIAHLVNNWELKQGGTLREAVLRAIPOLRGAYGTVMDSRHPDTLLAARSGSPLV	180	
QY	181	IGLGMGFASDQALLPVTRRFIFLEEGDIAEITRRSVNIPDKTGAEVKRODIESNLQ	240	
Db	181	IGLGMGFASDQALLPVTRRFIFLEEGDIAEITRRSVNIPDKTGAEVKRODIESNLQ	240	
QY	241	YDAGDKGIYCHYMQKEIYEOPNAIKNTLTGRISHGQVDLSLGPNADELLSKVEHQLA	300	
Db	241	YDAGDKGIYCHYMQKEIYEOPNAIKNTLTGRISHGQVDLSLGPNADELLSKVEHQLA	300	
QY	301	CGTSYNSGMVSRVWFESLAGIPCDVEIASSEFRYKSAVRNRSLMITLSQSGETADTLA	360	

```
Db 301 COTSNSGMVSRYPFESLAGIPCDVEIASFEFRYKSAVRNSLMTLSQSGETADTLACL 360
Qy 361 RLSKELGYLSLAI CNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQLTVLLMLVAKL 420
Db 361 RLSKELGYLSLAI CNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQLTVLLMLVAKL 420
Qy 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEFSDXKHALFLSRGDOYPIA 480
Db 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEFSDXKHALFLSRGDOYPIA 480
Qy 481 LEGALKLKEISYTHAEYAAGELKHGPLALIDADMPVIVVAPNNLEKLSKNIIEVRAR 540
Db 481 LEGALKLKEISYTHAEYAAGELKHGPLALIDADMPVIVVAPNNLEKLSKNIIEVRAR 540
Qy 541 GGQLYVFADQDAGFVSSDNMHIIEMPHVEEVIAPIFTYVPLQLLAYHVALIKGTDVDDQPR 600
Db 541 GGQLYVFADQDAGFVSSDNMHIIEMPHVEEVIAPIFTYVPLQLLAYHVALIKGTDVDDQPR 600
Qy 601 NLAKSVTVVE 609
Db 601 NLAKSVTVVE 609

RESULT 2
US-10-612-779-6
; Sequence 6, Application US/10612779
; Publication No. US20040091976A1
; GENERAL INFORMATION:
; APPLICANT: Deng, Ming-De
; APPLICANT: Angerer, J. David
; APPLICANT: Cyron, Don
; APPLICANT: Grund, Alan
; APPLICANT: Jerrell Jr., Thomas
; APPLICANT: Leanna, Candice
; APPLICANT: Mathre, Owen
; APPLICANT: Rosson, Reinhardt
; APPLICANT: Running, Jeff
; APPLICANT: Severson, Dave
; APPLICANT: Song, Linsheeng
; APPLICANT: Wassink, Sarah
; TITLE OF INVENTION: Processes and Materials for Production of Glucosamine and N-Acetyls
; FILE REFERENCE: 3161-18-2
; CURRENT APPLICATION NUMBER: US/10/612, 779
; CURRENT FILING DATE: 2003-07-01
; PRIOR APPLICATION NUMBER: 60/393, 348
; PRIOR FILING DATE: 2002-07-01
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-612-779-6

Query Match 100.0%; Score 3089; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 4.2e-252;
Matches 609; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MCGIVGAIQAQRDVAEILLEGRLRLEYRGYDSAGLAVVDTEGHMTLRLRLGKVQMLAQAAE 60
Db 1 MCGIVGAIQAQRDVAEILLEGRLRLEYRGYDSAGLAVVDTEGHMTLRLRLGKVQMLAQAAE 60
Qy 61 EHPHGGGTGIAHTRWATHGEPSEVNAHPVSHI VVVHNGIIEENHEPREEELKARGYTFV 120
Db 61 EHPHGGGTGIAHTRWATHGEPSEVNAHPVSHI VVVHNGIIEENHEPREEELKARGYTFV 120
Qy 121 SETDTEVIAHLVNWELKOGGTLREAVLRAI POLRGAYGTVMDSRHPDPTLLAARSGSPLV 180
Db 121 SETDTEVIAHLVNWELKOGGTLREAVLRAI POLRGAYGTVMDSRHPDPTLLAARSGSPLV 180
Qy 181 IGLGMGENFIASDQALLPVTRRRFIFLEEGDIAETTRRSVNI FDKTGAEVKQDIESNLQ 240
Db 181 IGLGMGENFIASDQALLPVTRRRFIFLEEGDIAETTRRSVNI FDKTGAEVKQDIESNLQ 240
```

```
Db 181 IGLGMGENFIASDQALLPVTRRRFIFLEEGDIAETTRRSVNI FDKTGAEVKQDIESNLQ 240
Qy 241 YDAGDKGIYCHYMKEIYEQPNAIKNTLTGRISHQQVDLSLPGNADELLSKVEHIQILA 300
Db 241 YDAGDKGIYCHYMKEIYEQPNAIKNTLTGRISHQQVDLSLPGNADELLSKVEHIQILA 300
Qy 301 COTSNSGMVSRYPFESLAGIPCDVEIASFEFRYKSAVRNSLMTLSQSGETADTLACL 360
Db 301 COTSNSGMVSRYPFESLAGIPCDVEIASFEFRYKSAVRNSLMTLSQSGETADTLACL 360
Qy 361 RLSKELGYLSLAI CNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQLTVLLMLVAKL 420
Db 361 RLSKELGYLSLAI CNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQLTVLLMLVAKL 420
Qy 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEFSDXKHALFLSRGDOYPIA 480
Db 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEFSDXKHALFLSRGDOYPIA 480
Qy 481 LEGALKLKEISYTHAEYAAGELKHGPLALIDADMPVIVVAPNNLEKLSKNIIEVRAR 540
Db 481 LEGALKLKEISYTHAEYAAGELKHGPLALIDADMPVIVVAPNNLEKLSKNIIEVRAR 540
Qy 541 GGQLYVFADQDAGFVSSDNMHIIEMPHVEEVIAPIFTYVPLQLLAYHVALIKGTDVDDQPR 600
Db 541 GGQLYVFADQDAGFVSSDNMHIIEMPHVEEVIAPIFTYVPLQLLAYHVALIKGTDVDDQPR 600
Qy 601 NLAKSVTVVE 609
Db 601 NLAKSVTVVE 609

RESULT 3
US-10-823-397-25
; Sequence 25, Application US/10823397
; Publication No. US20050042735A1
; GENERAL INFORMATION:
; APPLICANT: McMullin, Thomas
; APPLICANT: Ding, Ming-De
; APPLICANT: Grund, Alan
; TITLE OF INVENTION: Metabolic Engineering for Enhanced Production of Chitin and
; FILE REFERENCE: 3161-18-3
; CURRENT APPLICATION NUMBER: US/10/823, 397
; CURRENT FILING DATE: 2004-04-12
; PRIOR APPLICATION NUMBER: 60/462, 087
; PRIOR FILING DATE: 2003-04-11
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 25
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-823-397-25

Query Match 100.0%; Score 3089; DB 5; Length 609;
Best Local Similarity 100.0%; Pred. No. 4.2e-252;
Matches 609; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MCGIVGAIQAQRDVAEILLEGRLRLEYRGYDSAGLAVVDTEGHMTLRLRLGKVQMLAQAAE 60
Db 1 MCGIVGAIQAQRDVAEILLEGRLRLEYRGYDSAGLAVVDTEGHMTLRLRLGKVQMLAQAAE 60
Qy 61 EHPHGGGTGIAHTRWATHGEPSEVNAHPVSEHIVVVHNGIIEENHEPREEELKARGYTFV 120
Db 61 EHPHGGGTGIAHTRWATHGEPSEVNAHPVSEHIVVVHNGIIEENHEPREEELKARGYTFV 120
Qy 121 SETDTEVIAHLVNWELKOGGTLREAVLRAI POLRGAYGTVMDSRHPDPTLLAARSGSPLV 180
Db 121 SETDTEVIAHLVNWELKOGGTLREAVLRAI POLRGAYGTVMDSRHPDPTLLAARSGSPLV 180
Qy 181 IGLGMGENFIASDQALLPVTRRRFIFLEEGDIAETTRRSVNI FDKTGAEVKQDIESNLQ 240
Db 181 IGLGMGENFIASDQALLPVTRRRFIFLEEGDIAETTRRSVNI FDKTGAEVKQDIESNLQ 240
```

```
QY 241 YDAGDKGIYCHYMOKEIYEQPNAIKNTLTGRISHGOVDLSLGNADLLSKVEHIQILA 300
DB 241 YDAGDKGIYCHYMOKEIYEQPNAIKNTLTGRISHGOVDLSLGNADLLSKVEHIQILA 300
QY 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASEFRYKRSVAVRRNSLMTLSQSGETADTLA 360
DB 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASEFRYKRSVAVRRNSLMTLSQSGETADTLA 360
QY 361 RLSKELGYLGSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFTTQTLVLLMLVAKL 420
DB 361 RLSKELGYLGSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFTTQTLVLLMLVAKL 420
QY 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEFDSKHHALFLSRGDQYPIA 480
DB 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEFDSKHHALFLSRGDQYPIA 480
QY 481 LEGALKLKEISYTHAEAYAGELKHGFLALIDAMPVIVVAPNNLEKLEKSNIEVRAR 540
DB 481 LEGALKLKEISYTHAEAYAGELKHGFLALIDAMPVIVVAPNNLEKLEKSNIEVRAR 540
QY 541 GGOLYVPADQDAGFVSSDNHIIEMPHVEVIAPIFYTPVPLQLLAYHVALIKGTDVQDPR 600
DB 541 GGOLYVPADQDAGFVSSDNHIIEMPHVEVIAPIFYTPVPLQLLAYHVALIKGTDVQDPR 600
QY 601 NLAKSVTVE 609
DB 601 NLAKSVTVE 609

RESULT 4
US-10-024-460-28
; Sequence 28, Application US/10024460
; Publication No. US20030044939A1
; GENERAL INFORMATION:
; APPLICANT: Berry, Alan
; APPLICANT: Burlingame, Richard P.
; APPLICANT: Millis, James R.
; TITLE OF INVENTION: PROCESS AND MATERIALS FOR PRODUCTION OF GLUCOSAMINE
; FILE REFERENCE: 3161-18-C1
; CURRENT APPLICATION NUMBER: US/10/024,460
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/115,475
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/035,494
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-01-14
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-024-460-28

Query Match 99.4%; Score 3072; DB 4; Length 609;
Best Local Similarity 99.7%; Pred. No. 1.1e-250;
Matches 607; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MCGIVGAIQAORDVAEILLEGRLRLEYRGVDSAGLAVVDTEGHMTRRLRLGKVQMLAQAAE 60
DB 1 MCGIVGAIQAORDVAEILLEGRLRLEYRGVDSAGLAVVDTEGHMTRRLRLGKVQMLAQAAE 60
QY 61 EHPHCGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHHEPLREELKARGYTFV 120
DB 61 EHPHCGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHHEPLREELKARGYTFV 120
QY 121 SETDTEVIAHLVNWELKQGGLTLEAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGSPV 180
DB 121 SETDTEVIAHLVNWELKQGGLTLEAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGSPV 180
QY 181 IGLGMENFIASDQALLPVTRRFIFLEEGDIAEITRRSVNIFDKTGAEVKRODIESNLQ 240
DB 181 IGLGMENFIASDQALLPVTRRFIFLEEGDIAEITRRSVNIFDKTGAEVKRODIESNLQ 240
```

```
QY 241 YDAGDKGIYCHYMOKEIYEQPNAIKNTLTGRISHGOVDLSLGNADLLSKVEHIQILA 300
DB 241 YDAGDKGIYCHYMOKEIYEQPNAIKNTLTGRISHGOVDLSLGNADLLSKVEHIQILA 300
QY 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASEFRYKRSVAVRRNSLMTLSQSGETADTLA 360
DB 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASEFRYKRSVAVRRNSLMTLSQSGETADTLA 360
QY 361 RLSKELGYLGSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFTTQTLVLLMLVAKL 420
DB 361 RLSKELGYLGSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFTTQTLVLLMLVAKL 420
QY 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEFDSKHHALFLSRGDQYPIA 480
DB 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEFDSKHHALFLSRGDQYPIA 480
QY 481 LEGALKLKEISYTHAEAYAGELKHGFLALIDAMPVIVVAPNNLEKLEKSNIEVRAR 540
DB 481 LEGALKLKEISYTHAEAYAGELKHGFLALIDAMPVIVVAPNNLEKLEKSNIEVRAR 540
QY 541 GGOLYVPADQDAGFVSSDNHIIEMPHVEVIAPIFYTPVPLQLLAYHVALIKGTDVQDPR 600
DB 541 GGOLYVPADQDAGFVSSDNHIIEMPHVEVIAPIFYTPVPLQLLAYHVALIKGTDVQDPR 600
QY 601 NLAKSVTVE 609
DB 601 NLAKSVTVE 609

RESULT 5
US-10-024-460-31
; Sequence 31, Application US/10024460
; Publication No. US20030044939A1
; GENERAL INFORMATION:
; APPLICANT: Berry, Alan
; APPLICANT: Burlingame, Richard P.
; APPLICANT: Millis, James R.
; TITLE OF INVENTION: PROCESS AND MATERIALS FOR PRODUCTION OF GLUCOSAMINE
; FILE REFERENCE: 3161-18-C1
; CURRENT APPLICATION NUMBER: US/10/024,460
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/115,475
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/035,494
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-01-14
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-024-460-31

Query Match 99.4%; Score 3072; DB 4; Length 609;
Best Local Similarity 99.7%; Pred. No. 1.1e-250;
Matches 607; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MCGIVGAIQAORDVAEILLEGRLRLEYRGVDSAGLAVVDTEGHMTRRLRLGKVQMLAQAAE 60
DB 1 MCGIVGAIQAORDVAEILLEGRLRLEYRGVDSAGLAVVDTEGHMTRRLRLGKVQMLAQAAE 60
QY 61 EHPHCGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHHEPLREELKARGYTFV 120
DB 61 EHPHCGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHHEPLREELKARGYTFV 120
QY 121 SETDTEVIAHLVNWELKQGGLTLEAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGSPV 180
DB 121 SETDTEVIAHLVNWELKQGGLTLEAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGSPV 180
QY 181 IGLGMENFIASDQALLPVTRRFIFLEEGDIAEITRRSVNIFDKTGAEVKRODIESNLQ 240
DB 181 IGLGMENFIASDQALLPVTRRFIFLEEGDIAEITRRSVNIFDKTGAEVKRODIESNLQ 240
```

241	Qy	YDAGDKGIYCHYMKOEIYEQPNAIKNTUTGRISHGQVDLSLGPNADELLSKVEHIQILA	300
241	Db	YDAGDKGIYRHYMKOEIYEQPNAIKNTUTGRISHGQVDLSLGPNADELLSKVEHIQILA	300
301	Qy	CGTSYNSGMSRYWPFESLAGIPCDVEIASEFYRKSAVRRNSLMTLTSQSGETADTLAGL	360
301	Db	CGTSYNSGMSRYWPFESLAGIPCDVEIASEFYRKSAVRRNSLMTLTSQSGETADTLAGL	360
361	Qy	RLSKELGYLGSLATCNVPGSSILVRESDLALMTNAGTEIGVASTKAPTQTOLTVLMLVAKL	420
361	Db	RLSKELGYLGSLATCNVPGSSILVRESDLALMTNAGTEIGVASTKAPTQTOLTVLMLVAKL	420
421	Qy	SRLKGLDASIEHDIHVHGLQALPSRIEQMLSDQKRIEALAEFSDXHHALFLSRGQOYPIA	480
421	Db	SRLKGLDASIEHDIHVHGLQALPSRIEQMLSDQKRIEALAEFSDXHHALFLSRGQOYPIA	480
481	Qy	LEGALKLKEISYIHAEAYAAGELKHGPPLALTDADMPVIWVAPNNLELLEKLKSNIIEVRAR	540
481	Db	LEGALKLKEISYIHAEAYAAGELKHGPPLALTDADMPVIWVAPNNLELLEKLKSNIIEVRAR	540
541	Qy	GGQLYVFADQDAGFVSSDNMHIIEMPHVEEVIAPIFYTVPQLQLLAYHVALIKGTDDQDPR	600
541	Db	GGQLYVFADQDAGFVSSDNMHIIEMPHVEEVIAPIFYTVPQLQLLAYHVALIKGTDDQDPR	600
601	Qy	NLAKSVTVE	609
601	Db	NLAKSVTVE	609

RESULT 6  
US-10-612-779-10  
; Sequence 10, Application US/10612779  
; Publication No. US20040091976A1  
; GENERAL INFORMATION:  
; APPLICANT: Deng, Ming-De  
; APPLICANT: Angerer, J. David  
; APPLICANT: Cyron, Don  
; APPLICANT: Grund, Alan  
; APPLICANT: Jerrell Jr., Thomas  
; APPLICANT: Leanna, Candice  
; APPLICANT: Mathre, Owen  
; APPLICANT: Rosson, Reinhardt  
; APPLICANT: Running, Jeff  
; APPLICANT: Severson, Dave  
; APPLICANT: Song, Linsheng  
; APPLICANT: Wassink, Sarah  
; TITLE OF INVENTION: Process and Materials for Production of Glucosamine and N-Acetylle  
; TITLE OF INVENTION: Glucosamine  
; FILE REFERENCE: 3161-18-2  
; CURRENT APPLICATION NUMBER: US/10/612,779  
; CURRENT FILING DATE: 2003-07-01  
; PRIOR APPLICATION NUMBER: 60/393,348  
; PRIOR FILING DATE: 2002-07-01  
; NUMBER OF SEQ ID NOS: 137  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 10  
; LENGTH: 609  
; TYPE: PRT  
; ORGANISM: Escherichia coli  
US-10-612-779-10

Qy	121	SETDTEVIAHLVNWELKOGGTLRBAVRAIPQLRGAYGTVMDSRHPDTLLAARSGSPLV	180
Db	121	SETDTEVIAHLVNWELKOGGTLRBAVRAIPQLRGAYGTVMDSRHPDTLLAARSGSPLV	180
Qy	181	IGLGMGNFTASDQALLPVTRRFIFLEEGDIAELTRRSVNI FDKTGAEVKQDIESNIQ	240
Db	181	IGLGMGNFTASDQALLPVTRRFIFLEEGDIAELTRRSVNI FDKTGAEVKQDIESNIQ	240
Qy	241	YDAGDKGIYCHYMQKEIYEOPNAIKNTLTGRISHGQVDLSBELGPNADLLSKVEHIQILA	300
Db	241	YDAGDKGIYCHYMQKEIYEOPNAIKNTLTGRISHGQVDLSBELGPNADLLSKVEHIQILA	300
Qy	301	CGTSYNGMSRYWFESLAGIPCDVETASEFRYKSAVRNSLMITLSQSGETADTLAGL	360
Db	301	CGTSYNGMSRYWFESLAGIPCDVETASEFRYKSAVRNSLMITLSQSGETADTLAGL	360
Qy	361	RLSKELGYLGS LAICNYPGSSLVRESDLALMTNAGTEIGVASTKAFTTQLTVLMLVAKL	420
Db	361	RLSKELGYLGS LAICNYPGSSLVRESDLALMTNAGTEIGVASTKAFTTQLTVLMLVAKL	420
Qy	421	SRUKGLDASIEHDI VHGLQALPSRIEOMLSODKRIEALAEAFSDKHAFFLSRGQDYPYA	480
Db	421	SRUKGLDASIEHDI VHGLQALPSRIEOMLSODKRIEALAEAFSDKHAFFLSRGQDYPYA	480
Qy	481	LEGALKLUKEISYIHAEYAAAGELKHGPLALIDAMPVIVVAPNNLEKLKSNIEVRAR	540
Db	481	LEGALKLUKEISYIHAEYAAAGELKHGPLALIDAMPVIVVAPNNLEKLKSNIEVRAR	540
Qy	541	GGQLYVFADODAGFVSSDNMHHI IEMPHVEEVIADFTVTPQLLAIYHVALIKGTVDQDR	600
Db	541	GGQLYVFADODAGFVSSDNMHHI IEMPHVEEVIADFTVTPQLLAIYHVALIKGTVDQDR	600
Qy	601	NLAKSVTVE 609	
Db	601	NLAKSVTVE 609	

```

RESULT 7
US-10-612-779-12
; Sequence 12, Application US/10612779
; Publication No. US20040091976A1
; GENERAL INFORMATION:
; APPLICANT: Deng, Ming-De
; APPLICANT: Augerer, J. David
; APPLICANT: Cyron, Don
; APPLICANT: Grund, Alan
; APPLICANT: Jerrell Jr., Thomas
; APPLICANT: Leanna, Candice
; APPLICANT: Mathre, Owen
; APPLICANT: Rosson, Reinhardt
; APPLICANT: Running, Jeff
; APPLICANT: Severson, Dave
; APPLICANT: Song, Linsheng
; APPLICANT: Wassink, Sarah
; TITLE OF INVENTION: Process and Materials for Production of Glucosamine and N-Acetyls
; TITLE OF INVENTION: Glucosamine
; FILE REFERENCE: 3161-18-2
; CURRENT APPLICATION NUMBER: US/10/612,779
; CURRENT FILING DATE: 2003-07-01
; PRIOR APPLICATION NUMBER: 60/393,348
; PRIOR FILING DATE: 2002-07-01
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 609
; TYPE: prt
; ORGANISM: Escherichia coli
; US-10-612-779-12

```

QY 61 EPHLHGCGTGAHTRWATHGCEPSEVNAHPHYSEHVVVNHGIIENHEPLREELKARGYTFV 120

D6 61 EPHLHGCGTGAHTRWATHGCEPSEVNAHPHYSEHVVVNHGIIENHEPLREELKARGYTFV 120

Query Match 99.4%; Score 3072; DB 4; Length 609;  
Best Local Similarity 99.7%; Pred. No. 1.1e-250;  
Matches 607; Conservative 0; Mismatches 2; Indels



```
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/115,475
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/035,494
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-01-14
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-024-460-15

Query Match      99.3%; Score 3068; DB 4; Length 609;
Best Local Similarity 99.5%; Pred. No. 2.5e-250;
Matches 606; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MCGIVGATAQRDVAETILLEGRLRLEYRGYDSAGLAVVDTEGHMTLRLRLGKVQMLAQAAE 60
Db 1 MCGIVGATAQRDVAETILLEGRLRLEYRGYDSAGLAVVDTEGHMTLRLRLGKVQMLAQAAE 60
QY 61 EPHLGGTGTIAHTRWATHGEPSEVNAHPHVSEHIVVHNGIIEHHEPRLREELKARGYTFV 120
Db 61 EPHLGGTGTIAHTRWATHGEPSEVNAHPHVSEHIVVHNGIIEHHEPRLREELKARGYTFV 120
QY 121 SETDTEVIAHLVNWELKOGGTLREAVLRAIPQLRGAYGTVMDSRHPDTLLAARSGSPV 180
Db 121 SETDTEVIAHLVNWELKOGGTLREAVLRAIPQLRGAYGTVMDSRHPDTLLAARSGSPV 180
QY 181 IGLGNGENFIASDQALLPVTRRFIFLEEGDIAETRRSVNIFDKTGAEVKQDIESNLQ 240
Db 181 IGLGNGENFIASDQALLPVTRRFIFLEEGDIAETRRSVNIFDKTGAEVKQDIESNLQ 240
QY 241 YDAGDKGIYCHYMQKEIYEQPNAIKNTLTGRISHGQVDLSELGPNADLLSKVEHIQILA 300
Db 241 YDAGDKGIYCHYMQKEIYEQPNAIKNTLTGRISHGQVDLSELGPNADLLSKVEHIQILA 300
QY 301 CGTSYNSGMVSRYPFESLAGIPCDVEIASPEFRYKSAVRNLSMITLSQSGTADTLAGL 360
Db 301 CGTSYNSGMVSRYPFESLAGIPCDVEIASPEFRYKSAVRNLSMITLSQSGTADTLAGL 360
QY 361 RLSKELGYLGSIAICNVPSSILVRESDLALMTNAGTEIGVASTKAPTQTTLVLLMLVAKL 420
Db 361 RLSKELGYLGSIAICNVPSSILVRESDLALMTNAGTEIGVASTKAPTQTTLVLLMLVAKL 420
QY 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEAFSDKHHLFLSRGQYPIA 480
Db 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEAFSDKHHLFLSRGQYPIA 480
QY 481 LEGALKLKEISYIHAEAYAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEVRAR 540
Db 481 LEGALKLKEISYIHAEAYAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEVRAR 540
QY 541 GGQLVVFADQDAGFVSSDNMHIIEHPHVEEVIAPIFYTVPLQLLAYHVHVALIKGTVDQDPR 600
Db 541 GGQLVVFADQDAGFVSSDNMHIIEHPHVEEVIAPIFYTVPLQLLAYHVHVALIKGTVDQDPR 600
QY 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609

RESULT 10
US-10-369-493-23611
; Sequence 23611, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
```

```
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 23611
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-369-493-23611

Query Match      99.3%; Score 3068; DB 4; Length 609;
Best Local Similarity 99.5%; Pred. No. 2.5e-250;
Matches 606; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MCGIVGATAQRDVAETILLEGRLRLEYRGYDSAGLAVVDTEGHMTLRLRLGKVQMLAQAAE 60
Db 1 MCGIVGATAQRDVAETILLEGRLRLEYRGYDSAGLAVVDTEGHMTLRLRLGKVQMLAQAAE 60
QY 61 EPHLGGTGTIAHTRWATHGEPSEVNAHPHVSEHIVVHNGIIEHHEPRLREELKARGYTFV 120
Db 61 EPHLGGTGTIAHTRWATHGEPSEVNAHPHVSEHIVVHNGIIEHHEPRLREELKARGYTFV 120
QY 121 SETDTEVIAHLVNWELKOGGTLREAVLRAIPQLRGAYGTVMDSRHPDTLLAARSGSPV 180
Db 121 SETDTEVIAHLVNWELKOGGTLREAVLRAIPQLRGAYGTVMDSRHPDTLLAARSGSPV 180
QY 181 IGLGNGENFIASDQALLPVTRRFIFLEEGDIAETRRSVNIFDKTGAEVKQDIESNLQ 240
Db 181 IGLGNGENFIASDQALLPVTRRFIFLEEGDIAETRRSVNIFDKTGAEVKQDIESNLQ 240
QY 241 YDAGDKGIYCHYMQKEIYEQPNAIKNTLTGRISHGQVDLSELGPNADLLSKVEHIQILA 300
Db 241 YDAGDKGIYCHYMQKEIYEQPNAIKNTLTGRISHGQVDLSELGPNADLLSKVEHIQILA 300
QY 301 CGTSYNSGMVSRYPFESLAGIPCDVEIASPEFRYKSAVRNLSMITLSQSGTADTLAGL 360
Db 301 CGTSYNSGMVSRYPFESLAGIPCDVEIASPEFRYKSAVRNLSMITLSQSGTADTLAGL 360
QY 361 RLSKELGYLGSIAICNVPSSILVRESDLALMTNAGTEIGVASTKAPTQTTLVLLMLVAKL 420
Db 361 RLSKELGYLGSIAICNVPSSILVRESDLALMTNAGTEIGVASTKAPTQTTLVLLMLVAKL 420
QY 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEAFSDKHHLFLSRGQYPIA 480
Db 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEAFSDKHHLFLSRGQYPIA 480
QY 481 LEGALKLKEISYIHAEAYAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEVRAR 540
Db 481 LEGALKLKEISYIHAEAYAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEVRAR 540
QY 541 GGQLVVFADQDAGFVSSDNMHIIEHPHVEEVIAPIFYTVPLQLLAYHVHVALIKGTVDQDPR 600
Db 541 GGQLVVFADQDAGFVSSDNMHIIEHPHVEEVIAPIFYTVPLQLLAYHVHVALIKGTVDQDPR 600
QY 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609

RESULT 11
US-10-282-122A-56753
; Sequence 56753, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
```



APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John  
APPLICANT: Carr, Grant  
APPLICANT: Yamamoto, Robert  
APPLICANT: Forsyth, R.  
APPLICANT: Xu, H.  
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
FILE REFERENCE: ELITRA.034A  
CURRENT APPLICATION NUMBER: US/10/282,122A  
CURRENT FILING DATE: 2003-02-20  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/230,335  
PRIOR FILING DATE: 2000-09-06  
PRIOR APPLICATION NUMBER: 60/230,347  
PRIOR FILING DATE: 2000-09-09  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/267,636  
PRIOR FILING DATE: 2001-02-09  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 78614  
SOFTWARE: Patentin version 3.1  
SEQ ID NO 56753  
LENGTH: 609  
TYPE: PRT  
ORGANISM: Escherichia coli  
US-10-282-122A-56753

Query Match 99.3%; Score 3068; DB 4; Length 609;  
Best Local Similarity 99.5%; Pred. No. 2.5e-250;  
Matches 606; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MCGIVGAIQRDVAEILLGLRLRYGYDSAGLVAVDTTEGHMTRRLRLKQVOMLAQAAE 60  
DB 1 MCGIVGAIQRDVAEILLGLRLRYGYDSAGLVAVDAEGHMTLRLRLKQVOMLAQAAE 60

QY 61 EHPHGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHHEPLREELKARGYTFV 120  
DB 61 EHPHGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHHEPLREELKARGYTFV 120

QY 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDTLLAARSGSPV 180  
DB 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDTLLAARSGSPV 180

QY 181 IGLGMENFIASQDALLPVTRRFIFLEEGDIAEITRRSVNIIFDKTGAEVKRODIESNLQ 240  
DB 181 IGLGMENFIASQDALLPVTRRFIFLEEGDIAEITRRSVNIIFDKTGAEVKRODIESNLQ 240

QY 241 YDAGDKGIYCHYMKQEIYEQPNAIKNTLTGRI SHGOVDLSLGNADLLSKVEHIQILA 300  
DB 241 YDAGDKGIYCHYMKQEIYEQPNAIKNTLTGRI SHGOVDLSLGNADLLSKVEHIQILA 300

QY 301 CGTSYNSGMVSRVWFESLAGIPCDVEIFASEFRYKSAVRNLSMITLSQSGETADTLA 360  
DB 301 CGTSYNSGMVSRVWFESLAGIPCDVEIFASEFRYKSAVRNLSMITLSQSGETADTLA 360

QY 361 RLSKELGYLSLAICNVPGSSLVRESDLALMTWAGTEIGVASTKAFPTQTLTLLMLVAKL 420  
DB 361 RLSKELGYLSLAICNVPGSSLVRESDLALMTWAGTEIGVASTKAFPTQTLTLLMLVAKL 420

QY 421 SRLKGLDASIEHDIHVHGLQALPSRIEQLSODKRIEALAEFSDKHHALFLRGDQVPIA 480

DB 421 SRLKGLDASIEHDIHVHGLQALPSRIEQLSODKRIEALAEFSDKHHALFLRGDQVPIA 480  
QY 481 LEGALKLKEISYTHAEAYAGELKHCPALIDADMPVIVVAPNNELEKLSNIEVRAR 540  
DB 481 LEGALKLKEISYTHAEAYAGELKHCPALIDADMPVIVVAPNNELEKLSNIEVRAR 540  
QY 541 GGQLYVFADQDAGFVSSDNHIIEMPHVEVIAPIFYTVPLQLLAYHVALIKGTDDQPR 600  
DB 541 GGQLYVFADQDAGFVSSDNHIIEMPHVEVIAPIFYTVPLQLLAYHVALIKGTDDQPR 600  
QY 601 NLAKSVTVE 609  
DB 601 NLAKSVTVE 609

RESULT 12  
US-10-612-779-2  
; Sequence 2, Application US/10612779  
; Publication No. US20040091976A1  
; GENERAL INFORMATION:  
; APPLICANT: Deng, Ming-De  
; APPLICANT: Angerer, J. David  
; APPLICANT: Cyron, Don  
; APPLICANT: Grund, Alan  
; APPLICANT: Jerrell Jr., Thomas  
; APPLICANT: Leanna, Candice  
; APPLICANT: Mathre, Owen  
; APPLICANT: Rosson, Reinhardt  
; APPLICANT: Running, Jeff  
; APPLICANT: Severson, Dave  
; APPLICANT: Song, Linsheng  
; APPLICANT: Wassink, Sarah  
; TITLE OF INVENTION: Process and Materials for Production of Glucosamine and N-Acetyls  
; TITLE OF INVENTION: Glucosamine  
; FILE REFERENCE: 3161-18-2  
; CURRENT APPLICATION NUMBER: US/10/612,779  
; CURRENT FILING DATE: 2003-07-01  
; PRIOR APPLICATION NUMBER: 60/393,348  
; PRIOR FILING DATE: 2002-07-01  
; NUMBER OF SEQ ID NOS: 137  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 2  
; LENGTH: 609  
; TYPE: PRT  
; ORGANISM: Escherichia coli  
US-10-612-779-2

Query Match 99.3%; Score 3068; DB 4; Length 609;  
Best Local Similarity 99.5%; Pred. No. 2.5e-250;  
Matches 606; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MCGIVGAIQRDVAEILLGLRLRYGYDSAGLVAVDTTEGHMTRRLRLKQVOMLAQAAE 60  
DB 1 MCGIVGAIQRDVAEILLGLRLRYGYDSAGLVAVDAEGHMTLRLRLKQVOMLAQAAE 60

QY 61 EHPHGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHHEPLREELKARGYTFV 120  
DB 61 EHPHGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHHEPLREELKARGYTFV 120

QY 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDTLLAARSGSPV 180  
DB 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDTLLAARSGSPV 180

QY 181 IGLGMENFIASQDALLPVTRRFIFLEEGDIAEITRRSVNIIFDKTGAEVKRODIESNLQ 240  
DB 181 IGLGMENFIASQDALLPVTRRFIFLEEGDIAEITRRSVNIIFDKTGAEVKRODIESNLQ 240

QY 241 YDAGDKGIYCHYMKQEIYEQPNAIKNTLTGRI SHGOVDLSLGNADLLSKVEHIQILA 300  
DB 241 YDAGDKGIYCHYMKQEIYEQPNAIKNTLTGRI SHGOVDLSLGNADLLSKVEHIQILA 300

QY 301 CGTSYNSGMVSRVWFESLAGIPCDVEIFASEFRYKSAVRNLSMITLSQSGETADTLA 360

Db 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASPEFRYKSAVRNLSMITLSQSGETADTLA 360  
QY 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQLTVLLMLVAKL 420  
Db 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQLTVLLMLVAKL 420  
QY 421 SRLKGLDASIEHDIHVHGLQALPSRIEOMLSODKRIEALAEFSDKHAFILSRGDOYPTA 480  
Db 421 SRLKGLDASIEHDIHVHGLQALPSRIEOMLSODKRIEALAEFSDKHAFILSRGDOYPTA 480  
QY 481 LEGALKLKEISYTHAEAYAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEVRAR 540  
Db 481 LEGALKLKEISYTHAEAYAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEVRAR 540  
QY 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPIFTVPQLLAYHVALIKGTVDQPR 600  
Db 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPIFTVPQLLAYHVALIKGTVDQPR 600  
QY 601 NLAKSVTVE 609  
Db 601 NLAKSVTVE 609

RESULT 13  
US-10-823-397-23  
; Sequence 23, Application US/10823397  
; Publication No. US20050042735A1  
; GENERAL INFORMATION:  
; APPLICANT: McMullin, Thomas  
; APPLICANT: Ding, Ming-De  
; APPLICANT: Grund, Alan  
; TITLE OF INVENTION: Metabolic Engineering for Enhanced Production of Chitin and  
; FILE REFERENCE: 3161-18-3  
; CURRENT APPLICATION NUMBER: US/10/823,397  
; PRIOR FILING DATE: 2004-04-12  
; PRIOR APPLICATION NUMBER: 60/462,087  
; PRIOR FILING DATE: 2003-04-11  
; NUMBER OF SEQ ID NOS: 59  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 23  
; LENGTH: 609  
; TYPE: PRT  
; ORGANISM: Escherichia coli  
US-10-823-397-23

Query Match 99.3%; Score 3068; DB 5; Length 609;  
Best Local Similarity 99.5%; Pred. No. 2.5e-250;  
Matches 606; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 MCGIVGAIAQRDVAEILLEGRLRLRYRGYDSAGLAVVDTEGHMTLRLRLGKVQMLAQAAE 60  
Db 1 MCGIVGAIAQRDVAEILLEGRLRLRYRGYDSAGLAVVDTEGHMTLRLRLGKVQMLAQAAE 60  
QY 61 EHPHGGGTGIAHTRWATGEPSEVNAHPVSHIIVVHNGIIEHPELREELKARGYTFV 120  
Db 61 EHPHGGGTGIAHTRWATGEPSEVNAHPVSHIIVVHNGIIEHPELREELKARGYTFV 120  
QY 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPOLRGAYGTIVMDSRHPDTLLAARSGSPLV 180  
Db 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPOLRGAYGTIVMDSRHPDTLLAARSGSPLV 180  
QY 181 IGLGMGENFIASDQALLPVTRRFIFLEEGDIAETIRRSVNI FDKTGAEVKRDIESNLQ 240  
Db 181 IGLGMGENFIASDQALLPVTRRFIFLEEGDIAETIRRSVNI FDKTGAEVKRDIESNLQ 240  
QY 241 YDAGDKGYCHYMKQKEIYEOPNAIKNTLTGRISHGQVDLSELGNADLLSKVEHIQILA 300  
Db 241 YDAGDKGYCHYMKQKEIYEOPNAIKNTLTGRISHGQVDLSELGNADLLSKVEHIQILA 300  
QY 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASPEFRYKSAVRNLSMITLSQSGETADTLA 360  
Db 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASPEFRYKSAVRNLSMITLSQSGETADTLA 360

QY 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQLTVLLMLVAKL 420  
Db 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQLTVLLMLVAKL 420  
QY 421 SRLKGLDASIEHDIHVHGLQALPSRIEOMLSODKRIEALAEFSDKHAFILSRGDOYPTA 480  
Db 421 SRLKGLDASIEHDIHVHGLQALPSRIEOMLSODKRIEALAEFSDKHAFILSRGDOYPTA 480  
QY 481 LEGALKLKEISYTHAEAYAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEVRAR 540  
Db 481 LEGALKLKEISYTHAEAYAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEVRAR 540  
QY 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPIFTVPQLLAYHVALIKGTVDQPR 600  
Db 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPIFTVPQLLAYHVALIKGTVDQPR 600  
QY 601 NLAKSVTVE 609  
Db 601 NLAKSVTVE 609

RESULT 14  
US-10-024-460-25  
; Sequence 25, Application US/10024460  
; Publication No. US20030044939A1  
; GENERAL INFORMATION:  
; APPLICANT: Berry, Alan  
; APPLICANT: Burlingame, Richard P.  
; APPLICANT: Millis, James R.  
; TITLE OF INVENTION: PROCESS AND MATERIALS FOR PRODUCTION OF GLUCOSAMINE  
; FILE REFERENCE: 3161-18-C1  
; CURRENT APPLICATION NUMBER: US/10/024,460  
; CURRENT FILING DATE: 2001-12-17  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/115,475  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-15  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/035,494  
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-01-14  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 25  
; LENGTH: 609  
; TYPE: PRT  
; ORGANISM: Escherichia coli  
US-10-024-460-25

Query Match 99.1%; Score 3061; DB 4; Length 609;  
Best Local Similarity 99.3%; Pred. No. 9.7e-250;  
Matches 605; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 1 MCGIVGAIAQRDVAEILLEGRLRLRYRGYDSAGLAVVDTEGHMTLRLRLGKVQMLAQAAE 60  
Db 1 MCGIVGAIAQRDVAEILLEGRLRLRYRGYDSAGLAVVDTEGHMTLRLRLGKVQMLAQAAE 60  
QY 61 EHPHGGGTGIAHTRWATGEPSEVNAHPVSEHIVVHNGIIEHPELREELKARGYTFV 120  
Db 61 EHPHGGGTGIAHTRWATGEPSEVNAHPVSEHIVVHNGIIEHPELREELKARGYTFV 120  
QY 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPOLRGAYGTIVMDSRHPDTLLAARSGSPLV 180  
Db 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPOLRGAYGTIVMDSRHPDTLLAARSGSPLV 180  
QY 181 IGLGMGENFIASDQALLPVTRRFIFLEEGDIAETIRRSVNI FDKTGAEVKRDIESNLQ 240  
Db 181 IGLGMGENFIASDQALLPVTRRFIFLEEGDIAETIRRSVNI FDKTGAEVKRDIESNLQ 240  
QY 241 YDAGDKGYCHYMKQKEIYEOPNAIKNTLTGRISHGQVDLSELGNADLLSKVEHIQILA 300  
Db 241 YDAGDKGYCHYMKQKEIYEOPNAIKNTLTGRISHGQVDLSELGNADLLSKVEHIQILA 300  
QY 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASPEFRYKSAVRNLSMITLSQSGETADTLA 360  
Db 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASPEFRYKSAVRNLSMITLSQSGETADTLA 360

```
QY 361 RLSKELGYLGSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQLTVLLMLVAKL 420
DB 361 RLSKELGYLGSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQLTVLLMLVAKL 420
QY 421 SRLKGLDASTIEHDIHVHGLQALPSRIEQMLSDQKRIEALAEFSDKHHALFLSRGDQVPIA 480
DB 421 SRLKGLDASTIEHDIHVHGLQALPSRIEQMLSDQKRIEALAEFSDKHHAPFLGRGDQVPIA 480
QY 481 LEGALKLKEISYTHAEYAAAGELKHGFLALIDADMPVIVVAPNNELLEKLKSNIEVRAR 540
DB 481 LEGALKLKEISYTHAEYAAAGELKHGFLALIDADMPVIVVAPNNELLEKLKSNIEVRAR 540
QY 541 GGOLYVFADQDAGFVSSDNHIIEMPHVEVIAPIFTVPLQLLAYHVALIKGTDVDQPR 600
DB 541 GGOLYVFADQDAGFVSSDNHIIEMPHVEVIAPIFTVPLQLLAYHVALIKGTDVDQPR 600
QY 601 NLAKSVTVE 609
DB 601 NLAKSVTVE 609
```

## RESULT 15

US-10-612-779-8

; Sequence 8, Application US/10612779

; Publication No. US20040091976A1

; GENERAL INFORMATION:

; APPLICANT: Deng, Ming-De

; APPLICANT: Angerer, J. David

; APPLICANT: Cyron, Don

; APPLICANT: Grund, Alan

; APPLICANT: Jerrell Jr., Thomas

; APPLICANT: Leanna, Candice

; APPLICANT: Rosson, Owen

; APPLICANT: Rosson, Reinhardt

; APPLICANT: Running, Jeff

; APPLICANT: Severson, Dave

; APPLICANT: Song, Linsheng

; APPLICANT: Wassink, Sarah

; TITLE OF INVENTION: Process and Materials for Production of Glucosamine and N-Acetyls

; TITLE OF INVENTION: Glucosamine

; FILE REFERENCE: 3161-18-2

; CURRENT APPLICATION NUMBER: US/10/612,779

; PRIOR FILING DATE: 2003-07-01

; PRIOR APPLICATION NUMBER: 60/393,348

; PRIOR FILING DATE: 2002-07-01

; NUMBER OF SEQ ID NOS: 137

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 8

; LENGTH: 609

; TYPE: PRT

; ORGANISM: Escherichia coli

; US-10-612-779-8

```
Query Match          99.1%; Score 3061; DB 4; Length 609;
Best Local Similarity 99.3%; Pred. No. 9.7e-250;
Matches 605; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 1 MCGIVGAIARDVAEILLEGRLRLEVRGYDSAGLAVVDTEGHMTRLRRLGKVQMLAQAAE 60
DB 1 MCGIVGAIARDVAEILLEGRLRLEVRGYDSAGLAVVDTEGHMTRLRRLGKVQMLAQAAE 60
QY 61 EHPHGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIENHEPLREELKARGYTFV 120
DB 61 EHPHGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIENHEPLREELKARGYTFV 120
QY 121 SETDTEVIAHLNVNWLKQGTLREAVLRAIPOLRGAYGTVMDSRHPDTLLAARSGPLV 180
DB 121 SETDTEVIAHLNVNWLKQGTLREAVLRAIPOLRGAYGTVMDSRHPDTLLAARSGPLV 180
QY 181 IGLGMGENFIASDQALLPVTRRFIFLEEGDIAEITRRSVNIPDKTGAEVKRODIESNLQ 240
DB 181 IGLGMGENFIASDQALLPVTRRFIFLEEGDIAEITRRSVNIPDKTGAEVKRODIESNLQ 240
```

```
QY 241 YDAGDKGIYCHYMOKEIIYEQPNNAIKNTLTGRISHGOVDLSELGPNADELLSKVEHIOILA 300
DB 241 YDAGDKGIYCHYMOKEIIYEQPNNAIKNTLTGRISHGOVDLSELGPNADELLSKVEHIOILA 300
QY 301 CGTSYNSGMVSRVYWFESLAGIPCDVEIASSEFRYRKSAVRRNSLMTLSQSGETADTLAGL 360
DB 301 CGTSYNSGMVSRVYWFESLAGIPCDVEIASSEFRYRKSAVRRNSLMTLSQSGETADTLAGL 360
QY 361 RLSKELGYLGSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQLTVLLMLVAKL 420
DB 361 RLSKELGYLGSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQLTVLLMLVAKL 420
QY 421 SRLKGLDASTIEHDIHVHGLQALPSRIEQMLSDQKRIEALAEFSDKHHALFLSRGDQVPIA 480
DB 421 SRLKGLDASTIEHDIHVHGLQALPSRIEQMLSDQKRIEALAEFSDKHHAPFLGRGDQVPIA 480
QY 481 LEGALKLKEISYTHAEYAAAGELKHGFLALIDADMPVIVVAPNNELLEKLKSNIEVRAR 540
DB 481 LEGALKLKEISYTHAEYAAAGELKHGFLALIDADMPVIVVAPNNELLEKLKSNIEVRAR 540
QY 541 GGOLYVFADQDAGFVSSDNHIIEMPHVEVIAPIFTVPLQLLAYHVALIKGTDVDQPR 600
DB 541 GGOLYVFADQDAGFVSSDNHIIEMPHVEVIAPIFTVPLQLLAYHVALIKGTDVDQPR 600
QY 601 NLAKSVTVE 609
DB 601 NLAKSVTVE 609
```

Search completed: June 14, 2006, 15:34:24

Job time : 211.344 secs

**THIS PAGE IS BLANK**

GenCore version 5.1.9  
 Copyright (c) 1993 - 2006 Bioceleration Ltd.  
 OM protein - protein search, using sw model  
 Run on: June 14, 2006, 15:28:50 ; Search time 54.6562 Seconds  
 (without alignments)  
 1347.534 Million cell updates/sec  
 Title: US-10-612-779-30  
 Perfect score: 835  
 Sequence: 1 MSLPDGFYIRMEEGDLEQV.....VKFYKCGFSNAGVEMQIRK 159  
 Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5  
 Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : Published Applications AA Main:  
 1: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US07\_PUBCOMB.pcp.\*  
 2: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US08\_PUBCOMB.pcp.\*  
 3: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US09\_PUBCOMB.pcp.\*  
 4: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US10A\_PUBCOMB.pcp.\*  
 5: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US10B\_PUBCOMB.pcp.\*  
 6: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US11\_PUBCOMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	835	100.0	159	4	US-10-429-812-4
2	835	100.0	159	4	US-10-612-779-30
3	835	100.0	159	5	US-10-823-397-33
4	333	39.9	149	4	US-10-032-585-7301
5	328	39.3	149	4	US-10-612-779-32
6	328	39.3	149	5	US-10-823-397-35
7	275	32.9	170	4	US-10-128-714-8300
8	265.5	31.8	177	4	US-10-425-115-236099
9	261.5	31.3	149	4	US-10-612-779-34
10	234.5	28.1	174	4	US-10-424-599-159788
11	216.5	25.9	163	4	US-10-425-115-261686
12	216.5	25.9	174	4	US-10-425-114-61424
13	209	25.0	184	3	US-09-795-926-46
14	209	25.0	184	4	US-10-364-774-46
15	209	25.0	184	6	US-11-134-241-46
16	208	24.9	184	3	US-09-798-029-4
17	207	24.8	219	6	US-11-097-143-1509
18	204	24.4	184	4	US-10-092-900A-26
19	202.5	24.3	180	4	US-10-128-714-3300
20	202	24.2	184	4	US-10-092-900A-28
21	201.5	24.1	166	4	US-10-437-963-136699
22	199.5	23.9	165	4	US-10-437-963-116060
23	142	17.0	98	4	US-10-424-599-252530
24	116.5	14.0	157	6	US-11-045-004-2748
25	116	13.9	160	4	US-10-156-761-12507
26	114	13.7	133	3	US-09-798-029-25
27	112	13.4	140	4	US-10-282-122A-51716

28 110 13.2 179 3 US-09-738-626-4615 Sequence 4615, Ap  
 29 110 13.2 179 5 US-10-703-7998-20 Sequence 20, Appl  
 30 106 12.7 185 6 US-11-045-004-264 Sequence 264, Appl  
 31 105.5 12.6 159 4 US-10-433-256-9 Sequence 9, Appl  
 32 101 12.1 168 6 US-10-282-122A-60831 Sequence 60831, A  
 33 101 12.1 168 6 US-11-045-004-1735 Sequence 1735, Ap  
 34 101 12.1 170 6 US-11-124-367A-459 Sequence 459, Appl  
 35 98 11.7 23 3 US-09-927-734C-5 Sequence 5, Appl  
 36 97.5 11.7 147 5 US-10-501-282-2468 Sequence 2468, Ap  
 37 97.5 11.7 197 5 US-10-501-282-2470 Sequence 2470, Ap  
 38 97 11.6 218 4 US-10-425-115-339875 Sequence 339875,  
 39 96 11.5 254 4 US-10-437-963-160771 Sequence 160771,  
 40 96 11.5 257 4 US-10-424-599-273495 Sequence 273495,  
 41 96 11.5 263 4 US-10-425-114-50819 Sequence 50819, A  
 42 95 11.4 104 4 US-10-425-115-275860 Sequence 275860,  
 43 95 11.4 371 4 US-10-425-115-315609 Sequence 315609,  
 44 94.5 11.3 151 6 US-11-045-004-1609 Sequence 1609, Ap  
 45 93 11.1 192 4 US-10-424-599-147429 Sequence 147429,

## ALIGNMENTS

RESULT 1  
 US-10-429-812-4  
 ; Sequence 4, Application US/10429812  
 ; Publication No. US20040003432A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Obukowicz, Mark  
 ; TITLE OF INVENTION: THE PRODUCTION OF HEXOSAMINES AND USES THEREFORE  
 ; FILE REFERENCE: 18438/09004  
 ; CURRENT APPLICATION NUMBER: US/10/429,812  
 ; CURRENT FILING DATE: 2003-05-05  
 ; NUMBER OF SEQ ID NOS: 14  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 4  
 ; LENGTH: 159  
 ; TYPE: PRT  
 ; ORGANISM: Saccharomyces cerevisiae  
 US-10-429-812-4

Query Match 100.0%; Score 835; DB 4; Length 159;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-82;  
 Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MSLPDGFYIRMEEGDLEQVTTKVLTVGTITPESFSLIKYWNATVWVNDKIM 60  
 DB 1 MSLPDGFYIRMEEGDLEQVTTKVLTVGTITPESFSLIKYWNATVWVNDKIM 60  
 QY 61 QYNPWVVDKRTETVAATGNIIIERKIHGELGCHIEDIAVNSKYOGGKLLIDQLV 120  
 DB 61 QYNPWVVDKRTETVAATGNIIIERKIHGELGCHIEDIAVNSKYOGGKLLIDQLV 120  
 QY 121 TIGFDYGCYKIIIDCDCKNKVFEKCGFSNAGVEMQIRK 159  
 DB 121 TIGFDYGCYKIIIDCDCKNKVFEKCGFSNAGVEMQIRK 159

RESULT 2  
 US-10-612-779-30  
 ; Sequence 30, Application US/10612779  
 ; Publication No. US20040091976A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Deng, Ming-De  
 ; APPLICANT: Angerer, J. David  
 ; APPLICANT: Cyron, Don  
 ; APPLICANT: Grund, Alan  
 ; APPLICANT: Jerrell Jr., Thomas  
 ; APPLICANT: Leanna, Candice  
 ; APPLICANT: Mathre, Owen  
 ; APPLICANT: Rosson, Reinhardt  
 ; APPLICANT: Running, Jeff  
 ; APPLICANT: Severson, Dave

```
; APPLICANT: Song, Linsheng
; APPLICANT: Wassink, Sarah
; TITLE OF INVENTION: Process and Materials for Production of Glucosamine and N-Acetyls
; TITLE OF INVENTION: Glucosamine
; FILE REFERENCE: 3161-18-2
; CURRENT APPLICATION NUMBER: US/10/612,779
; CURRENT FILING DATE: 2003-07-01
; PRIOR APPLICATION NUMBER: 60/393,348
; PRIOR FILING DATE: 2002-07-01
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-612-779-30

Query Match          100.0%; Score 835; DB 4; Length 159;
Best Local Similarity 100.0%; Pred. No. 1.1e-82;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSLPDGFYIRMEGDLQVETLTKVLTVTGTTTPESFSKLIKYWNEATVWNNDEKKIM 60
Db 1 MSLPDGFYIRMEGDLQVETLTKVLTVTGTTTPESFSKLIKYWNEATVWNNDEKKIM 60

Qy 61 QYNPMVVDKRTETVAATGNIIIRKIIHELGLGCHIEDIAVNSKYQGGLGLLLIDQLV 120
Db 61 QYNPMVVDKRTETVAATGNIIIRKIIHELGLGCHIEDIAVNSKYQGGLGLLLIDQLV 120

Qy 121 TIGFDYGCYKIILDCDEKNVKFYKCGFSNAGVEMQIRK 159
Db 121 TIGFDYGCYKIILDCDEKNVKFYKCGFSNAGVEMQIRK 159

RESULT 3
US-10-823-397-33
; Sequence 33, Application US/10823397
; Publication No. US20050042735A1
; GENERAL INFORMATION:
; APPLICANT: McMullin, Thomas
; APPLICANT: Ding, Ming-De
; APPLICANT: Grund, Alan
; TITLE OF INVENTION: Metabolic Engineering for Enhanced Production of Chitin and
; TITLE OF INVENTION: Chitosan in Microorganisms
; FILE REFERENCE: 3161-18-3
; CURRENT APPLICATION NUMBER: US/10/823,397
; CURRENT FILING DATE: 2004-04-12
; PRIOR APPLICATION NUMBER: 60/462,087
; PRIOR FILING DATE: 2003-04-11
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 33
; LENGTH: 159
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-823-397-33

Query Match          100.0%; Score 835; DB 5; Length 159;
Best Local Similarity 100.0%; Pred. No. 1.1e-82;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSLPDGFYIRMEGDLQVETLTKVLTVTGTTTPESFSKLIKYWNEATVWNNDEKKIM 60
Db 1 MSLPDGFYIRMEGDLQVETLTKVLTVTGTTTPESFSKLIKYWNEATVWNNDEKKIM 60

Qy 61 QYNPMVVDKRTETVAATGNIIIRKIIHELGLGCHIEDIAVNSKYQGGLGLLLIDQLV 120
Db 61 QYNPMVVDKRTETVAATGNIIIRKIIHELGLGCHIEDIAVNSKYQGGLGLLLIDQLV 120

Qy 121 TIGFDYGCYKIILDCDEKNVKFYKCGFSNAGVEMQIRK 159
Db 121 TIGFDYGCYKIILDCDEKNVKFYKCGFSNAGVEMQIRK 159
```

## RESULT 4

```
US-10-032-585-7301
; Sequence 7301, Application US/10032585
; Publication No. US20030180953A1
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: Terry, Roemer D.
```

```
; APPLICANT: Bo, Jiang
```

```
; APPLICANT: Charles, Boone
```

```
; APPLICANT: Howard, Bussey
```

```
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
```

```
; FILE REFERENCE: 10182-005-999
```

```
; CURRENT APPLICATION NUMBER: US/10/032,585
```

```
; CURRENT FILING DATE: 2001-12-20
```

```
; NUMBER OF SEQ ID NOS: 8000
```

```
; SOFTWARE: PatentIn version 3.1
```

```
; SEQ ID NO 7301
```

```
; LENGTH: 149
```

```
; TYPE: PRT
```

```
; ORGANISM: Candida albicans
```

```
US-10-032-585-7301
```

```
Query Match          39.9%; Score 333; DB 4; Length 149;
```

```
Best Local Similarity 44.0%; Pred. No. 6.2e-28;
```

```
Matches 70; Conservative 26; Mismatches 51; Indels 12; Gaps 3;
```

```
Qy 1 MSLPDGFYIRMEGDLQVETLTKVLTVTGTTTPESFSKLIKYWNEATVWNNDEKKI 59
```

```
Db 1 MSLPDGFYIRMEGDLQVETLTKVLTVTGTTTPESFSKLIKYWNEATVWNNDEKKI 59
```

```
Qy 60 MQYNPMVVDKRTETVAATGNIIIRKIIHELGLGCHIEDIAVNSKYQGGLGLLLIDQL 119
```

```
Db 51 SIYHPYVITN-ASGIVVATGMLFVEKKLIHECGKVGHIEDISVAKSEQKGLGYLVTS 109
```

```
Qy 120 VTIGFDYGCYKIILDCDEKNVKFYKCGFSNAGVEMQIR 158
```

```
Db 110 TKVAQENDCYKVIILDCSPENVGFYKCGYKDGGMVCMR 148
```

## RESULT 5

```
US-10-612-779-32
```

```
; Sequence 32, Application US/10612779
```

```
; Publication No. US20040091976A1
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: Deng, Ming-De
```

```
; APPLICANT: Angerer, J. David
```

```
; APPLICANT: Cyron, Don
```

```
; APPLICANT: Grund, Alan
```

```
; APPLICANT: Jerrell Jr., Thomas
```

```
; APPLICANT: Leanna, Candice
```

```
; APPLICANT: Mathre, Owen
```

```
; APPLICANT: Rosson, Reinhardt
```

```
; APPLICANT: Running, Jeff
```

```
; APPLICANT: Severson, Dave
```

```
; APPLICANT: Song, Linsheng
```

```
; APPLICANT: Wassink, Sarah
```

```
; TITLE OF INVENTION: Process and Materials for Production of Glucosamine
```

```
; FILE REFERENCE: 3161-18-2
```

```
; CURRENT APPLICATION NUMBER: US/10/612,779
```

```
; CURRENT FILING DATE: 2003-07-01
```

```
; PRIOR APPLICATION NUMBER: 60/393,348
```

```
; PRIOR FILING DATE: 2002-07-01
```

```
; NUMBER OF SEQ ID NOS: 137
```

```
; SOFTWARE: PatentIn version 3.1
```

```
; SEQ ID NO 32
```

```
; LENGTH: 149
```

```
; TYPE: PRT
```

```
; ORGANISM: Candida albicans
```

```
US-10-612-779-32
```

```
Query Match          39.3%; Score 328; DB 4; Length 149;
```

```
Best Local Similarity 43.4%; Pred. No. 2.2e-27;
```

	Matches	69;	Conservative	26;	Mismatches	52;	Indels	12;	Gaps	3;
Qy	1	MSLPDGVIRRMEEGDLE-QVTETLKVLTITVGTTPSFSKLIKYNEATVMNDNEDKKI	59							
		: : :   :                 : :   :   :								
Db	1	MLPQGTYFRKLKLTIDYDNQYLETKVLTTVGESIKEDFTELYNNHWSLP-----	50							
		: : :   :                 : :   :   :								
Qy	60	MQYNPMVIVDKRTTVAQTGNIIERKKIIHELGLCGHIEDIAVNSKYGQGGGLKLLIQDL	119							
		: : :   :                 : :   :   :								
Db	51	SIYHPYVIITN-ASGIIVATGMFLFVEKKLIHECGKVGHIEDISVAKSEGGKLYGYLVTSL	109							
		: : :   :                 : :   :   :								
Qy	120	VTIQFDYGCYKIILDDCKENKVFYEKCGFSNAGVEMQIR	158							
		: : :                     : : :								
Db	110	TKVAOENDCYKVIDLCSPENNVGFYEKCYKDGGYEMVCR	148							
		: : :                     : : :								

```

RESULT 6
US-10-823-397-35
; Sequence 35, Application US/10823397
; Publication No. US20050042735A1
; GENERAL INFORMATION:
; APPLICANT: McMullin, Thomas
; APPLICANT: Ding, Ming-De
; APPLICANT: Grund, Alan
; TITLE OF INVENTION: Metabolic Engineering for Enhanced Production of Chitin and
; TITLE OF INVENTION: Chitosan in Microorganisms
; FILE REFERENCE: 3161-18-3
; CURRENT APPLICATION NUMBER: US/10/823,397
; CURRENT FILING DATE: 2004-04-12
; PRIOR APPLICATION NUMBER: 60/462,087
; PRIOR FILING DATE: 2003-04-11
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 35
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-823-397-35

```

Query Match	39.3%	Score 328;	DB 5;	Length 149;
Best Local Similarity	43.4%;	Pred. No. 2.2e-27;		
Matches	69;	Conservative 26;	Mismatches 52;	Indels 12; Gaps 3;
Qy	1	MSLPDGFVIRMEBGDL E-QVTETKLVLTGTTITPESFKLIKYMNEATVNNDKKI	59	
Db	1	MLPLPGGYTFRKLKLTVDNOYLETKVLTVTGEISKEDFTELYNHWSLP-----	50	
Qy	60	MQYNPMIVDKRTETAQTNIIERKIITHELGCGHIEDIAVNSKYQGOGGLKLIDQL	119	
Db	51	SIYHPYITN-ASGIIVATGMLFVEKLLIHCCGVGHIEDISVAKSQQKLGYYLVTSL	109	
Qy	120	VTI GF D Y C Y K I I L D C D E K N V K F Y E K C G F S N A G V E M Q I R	158	
Db	110	T K V A O E N D C Y K V I L D C S P E N V G F Y E K C G Y K D G G V E M V C R	148	

```

RESULT 7
US-10-128-714-8300
; Sequence 8300, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wenqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroshekin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; TITLE OF INVENTION: Methods of Use
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23

```

```

, PRIOR APPLICATION NUMBER: US 60/287,066
, PRIOR FILING DATE: 2001-04-27
, PRIOR APPLICATION NUMBER: US 60/295,890
, PRIOR FILING DATE: 2001-06-05
, PRIOR APPLICATION NUMBER: US 60/303,899
, PRIOR FILING DATE: 2001-07-09
, PRIOR APPLICATION NUMBER: US 60/316,362
, PRIOR FILING DATE: 2001-08-31
, NUMBER OF SEQ ID NOS: 8603
, SOFTWARE: PatentIn version 3.1
, SEQ ID NO 8300
, LENGTH: 190
, TYPE: PRT
, ORGANISM: Aspergillus fumigatus
US-10-128-714-8300

```

Query Match	32.9%	Score 275;	DB 4;	Length 190;
Best Local Similarity	41.3%;	Pred. No. 1.8e-21;		
Matches	64;	Conservative 19;	Mismatches 58;	Indels 14; Gaps 5;
Qy	3	LPDGFYIRRMEEGLEQ-VTETLKVLTVTGVTITPESFSKLIKYNNEATVW-NDNEDKKIM	60	
Db	44	LPADYITRPLCRSDYKKGYLVDLRLVTVGDIINEEQ-----WNSRYEWIARSD----	92	
Qy	61	QYNPMWVDRKTRTVAATGNIIIBRKIIHELGLGCHIEDIAVNSKYQGQGLKLIDQIV	120	
Db	93	EYILLVVCDBEGR-IVTGSLSVVERKPIHSLGMVGHIEDIAVEKQGQGLGLRIIQALD	151	
Qy	121	TIGFDYGCYKIILDCDEKNVKFYKCKGFSNAGVEM	155	
Db	152	YVASKVGICYKIIIDCSANEGFYIKCGFKRAGLEM	186	

```

RESULT 8
US-10-425-115-236099
; Sequence 236099, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 236099
; LENGTH: 177
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(177)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MFT4577_146905C.1.pep
US-10-425-115-236099

```

Query Match	31.8%	Score 265.5;	DB 4;	Length 177;
Best Local Similarity	41.1%	Pred. No. 1.8e-20;		
Matches	65;	Conservative 21;	Mismatches 57;	Indels 15; Gaps 5;
Qy	2	SLPDGPIRMEEGDLEQ-VTEILKVLVTITGGTITPESFSLIKIYKNEATVWNDNEKKIM	60	
Db	19	ALPEGYTLRALRKSDFNFGFLDCLRLVTITGGDITEADFKV-QYDDMAAAGS-----	68	
Qy	61	QYNPMWIVD--KRTETVAATGNIIIRKIIITHELGLCGHIE-DIAVNSKYOGQGGKLLID	117	
Db	69	-YIIIIIDTSRKENPVVGTGALITERKFTIHSIGAVGHIEXDIAVAKDQGGKLLGRLIIQ	127	
Qy	118	QLVTIGTDPDGYCKIIILDCDEKNKVFYKCGFSNAGVEM	155	





```

; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 61424
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3059-059-G11_FLI pep
US-10-425-114-61424

Query Match      25.0%; Score 216.5; DB 4; Length 174;
Best Local Similarity 33.8%; Pred. No. 3.9e-15;
Matches 52; Conservative 28; Mismatches 65; Indels 9; Gaps 3;

Qy 5 DGFYIRMEEGDLEQ-VTETLKVLTVTGTTTPEFSKLIKYNWEATVWMDNEDKKIMQYN 63
Db 27 DTTHIRLECSHERGFVALLSQLSPDOLTTSVFATRF-----AELAAQGDHILVAE 81

Qy 64 PMWIVDKRTETVAATGNIIIERKIIHELGLCGHIEDIAVNSKYQGGLKLLIDQLVTVIG 123
Db 82 DPSASDR---ILATCLFVERKFLRGCKGVGHVVDVDAARGGLGLRIVRLVEIS 138

Qy 124 FDYGCYKIILDCDEKNVKFYKCGFSNAGVEMQI 157
Db 139 RDAGCYKILDTPELRAYYAKGFGVEKGVQNAV 172

RESULT 13
US-09-795-926-46
; Sequence 46, Application US/09795926
; Patent No. US20020098486A1
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; FILE REFERENCE: LEX-0144-USA
; CURRENT APPLICATION NUMBER: US/09/795,926
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 184
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-795-926-46

Query Match      25.0%; Score 209; DB 3; Length 184;
Best Local Similarity 32.9%; Pred. No. 2.8e-14;
Matches 48; Conservative 25; Mismatches 61; Indels 12; Gaps 2;

Qy 5 DGFYIRMEEGDLEQ-VTETLKVLTVTGTTTPEFSKLIKYNWEATVWMDNEDKKIMQYN 63
Db 37 EGLVLRPLCTADLNRGFFKVLGOLTGTGVVSPQFMKSFEHM-----KKSQDY 85

Qy 64 PMWIVDKRTETVAATGNIIIERKIIHELGLCGHIEDIAVNSKYQGGLKLLIDQLVTVIG 123
Db 86 VTVVEDVTLGQIVATATLIIIEHKFIHSCAKRGVEDVWVSDCEGRQGLKLLSTLTLLS 145

Qy 124 FDYGCYKIILDCDEKNVKFYKCGFS 149
Db 146 KKLNCYKITLECLPQNVGFYKVGYYT 171

RESULT 14
US-10-364-774-46
; Sequence 46, Application US/10364774
; Publication No. US20030144497A1
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; FILE REFERENCE: LEX-0144-USA
; CURRENT APPLICATION NUMBER: US/10/364,774
; CURRENT FILING DATE: 2003-02-11
; PRIOR APPLICATION NUMBER: US/09/795,926
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 184
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-364-774-46

Query Match      25.0%; Score 209; DB 4; Length 184;
Best Local Similarity 32.9%; Pred. No. 2.8e-14;
Matches 48; Conservative 25; Mismatches 61; Indels 12; Gaps 2;

Qy 5 DGFYIRMEEGDLEQ-VTETLKVLTVTGTTTPEFSKLIKYNWEATVWMDNEDKKIMQYN 63
Db 37 EGLVLRPLCTADLNRGFFKVLGOLTGTGVVSPQFMKSFEHM-----KKSQDY 85

Qy 64 PMWIVDKRTETVAATGNIIIERKIIHELGLCGHIEDIAVNSKYQGGLKLLIDQLVTVIG 123
Db 86 VTVVEDVTLGQIVATATLIIIEHKFIHSCAKRGVEDVWVSDCEGRQGLKLLSTLTLLS 145

Qy 124 FDYGCYKIILDCDEKNVKFYKCGFS 149
Db 146 KKLNCYKITLECLPQNVGFYKVGYYT 171

RESULT 15
US-11-134-241-46
; Sequence 46, Application US/11134241
; Publication No. US20050287568A1
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn

```



Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	199.5	23.9	165	6	US-10-449-902-34519	Sequence 34519, A
2	111	13.3	166	6	US-10-471-571A-3964	Sequence 3964, Ap
3	98	11.7	257	6	US-10-953-349-1260	Sequence 1260, Ap
4	96	11.5	254	6	US-10-449-902-29790	Sequence 29790, A
5	92	11.0	159	6	US-10-953-349-2743	Sequence 2743, Ap
6	92	11.0	223	6	US-10-953-349-2742	Sequence 2742, Ap
7	92	11.0	237	6	US-10-953-349-2741	Sequence 2741, Ap
8	88.5	10.6	149	6	US-10-471-571A-5116	Sequence 5116, Ap
9	83.5	10.0	139	6	US-10-370-959-55	Sequence 55, Appl
10	83.5	10.0	286	6	US-10-471-571A-2210	Sequence 2210, Ap
11	79.5	9.5	154	6	US-10-471-571A-260	Sequence 260, App
12	79.5	9.5	492	6	US-10-449-902-42958	Sequence 42958, A
13	79.5	9.5	492	6	US-10-449-902-52839	Sequence 52839, A
14	79	9.5	133	6	US-10-471-571A-4442	Sequence 4442, Ap
15	77	9.2	435	6	US-10-953-349-10209	Sequence 10209, A
16	77	9.2	603	6	US-10-953-349-10208	Sequence 10208, A
17	77	9.2	609	6	US-10-953-349-10207	Sequence 10207, A
18	75.5	9.0	158	6	US-10-471-571A-3244	Sequence 3244, Ap
19	75	9.0	317	6	US-10-449-902-31187	Sequence 31187, A
20	74.5	8.9	271	6	US-10-449-902-50947	Sequence 50947, A
21	74.5	8.9	466	6	US-10-471-571A-5066	Sequence 5066, Ap
22	74	8.9	711	6	US-10-471-571A-2458	Sequence 2458, Ap
23	73.5	8.8	183	6	US-10-471-571A-4252	Sequence 4252, Ap
24	73	8.7	251	6	US-10-953-349-26429	Sequence 26429, A
25	73	8.7	290	6	US-10-953-349-26428	Sequence 26428, A

```
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927W0
; CURRENT APPLICATION NUMBER: US/10/471,571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 3964
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (1)..(166)
; OTHER INFORMATION: conserved hypothetical
US-10-471-571A-3964

Query Match      13.3%; Score 111; DB 6; Length 166;
Best Local Similarity 26.5%; Pred. No. 0.00063;
Matches 41; Conservative 26; Mismatches 64; Indels 24; Gaps 6;

Qy 9 IRRMEGDLQVETLKVLTVG-----TTPSFSLIKYKNEATVWNNEDKKIMQYN 63
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3 LRKIVMDLQII-----ALENIGFSPBEAATREALKLRIEQITFTIVAENKDEVIGYN 58
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 64 -PMW----IVDKRTTVAATGNTIIERKIIHELGLCGHIEDIAVNSKYQGGLGKLLIDQ 118
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 59 GPVIKERYISDLEFKNVSTNNS---EGGYISVLGL-----VVAPNYQGGIAGRLNY 108
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 119 LVITGFYGYCKIILDCDEKNVFPYKGFSGNAGV 153
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 109 FETLAKNHRHGVTLTCRESLISFYKGYRNEGV 143
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 3
US-10-953-349-1260
; Sequence 1260, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1260
; LENGTH: 257
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-1260

Query Match      11.7%; Score 98; DB 6; Length 257;
Best Local Similarity 42.3%; Pred. No. 0.02;
Matches 22; Conservative 8; Mismatches 22; Indels 0; Gaps 0;

Qy 97 IEDIAVNSKYQGGLGKLLIDQLVTIGFDYGCYKIIILDCDEKNVFKYKCGF 148
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 191 IWDVLVDPESYQGGLGKALVEKLVALLQRLDIGNISLPADSQVVDYFNGLGF 242
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 4
US-10-449-902-29790
; Sequence 29790, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
```

```
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29790
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-29790

Query Match      11.5%; Score 96; DB 6; Length 254;
Best Local Similarity 40.4%; Pred. No. 0.031;
Matches 21; Conservative 8; Mismatches 23; Indels 0; Gaps 0;

Qy 97 IEDIAVNSKYQGGLGKLLIDQLVTIGFDYGCYKIIILDCDEKNVFKYKCGF 148
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 187 IWDVLVDPESYQGGLGKALMEKVIITLLQRLDISNITLFADNKVVDYFNGLGF 238
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 5
US-10-953-349-2743
; Sequence 2743, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2743
; LENGTH: 159
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-2743

Query Match      11.0%; Score 92; DB 6; Length 159;
Best Local Similarity 27.0%; Pred. No. 0.042;
Matches 20; Conservative 16; Mismatches 38; Indels 0; Gaps 0;

Qy 77 ATGNIIEIKIHELGLCGHIEDIAVNSKYQGGLGKLLIDQLVTIGFDYGCYKIIILDCD 136
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 70 SNGQLVGFGRVSDYGLTASIHDLMLVLPQLQRMGIGKLIIVNRIVRLTTSRDIYDIAALCF 129
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 137 EKNVFKYKCGFSN 150
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 130 EDERPFFKACGFGD 143
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 6
US-10-953-349-2742
; Sequence 2742, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2742
; LENGTH: 223
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
```

## US-10-953-349-2742

Query Match 11.0%; Score 92; DB 6; Length 223;  
Best Local Similarity 27.0%; Pred. No. 0.064; Mismatches 38; Indels 0; Gaps 0;  
Matches 20; Conservative 16; Mismatches 38; Indels 0; Gaps 0;  
QY 77 ATGNIIRKIIHELGLCHIEDIAVNSKYQGGLKLLIDQLVTIGFDYGCYKILDCD 136  
DB 134 SNGQLVGFRAYSDELGTASIDHLMVLSPLQRMGIGKLVNRIVLLTSRDIYDIALCF 193  
QY 137 EKNVKFYKCGFSN 150  
DB 194 EDERPFFKACGFGD 207

## RESULT 7

US-10-953-349-2741  
; Sequence 2741, Application US/10953349  
; Publication No. US20060107345A1  
; GENERAL INFORMATION:  
; APPLICANT: ALEXANDROV, Nickolai et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; FILE REFERENCE: 2750-1579PUS2  
; CURRENT APPLICATION NUMBER: US/10/953,349  
; CURRENT FILING DATE: 2004-09-30  
; NUMBER OF SEQ ID NOS: 40252  
; SOFTWARE: Patent in version 3.3  
; SEQ ID NO 2741  
; LENGTH: 237  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-10-953-349-2741

Query Match 11.0%; Score 92; DB 6; Length 237;  
Best Local Similarity 27.0%; Pred. No. 0.069; Mismatches 38; Indels 0; Gaps 0;  
Matches 20; Conservative 16; Mismatches 38; Indels 0; Gaps 0;  
QY 77 ATGNIIRKIIHELGLCHIEDIAVNSKYQGGLKLLIDQLVTIGFDYGCYKILDCD 136  
DB 148 SNGQLVGFRAYSDELGTASIDHLMVLSPLQRMGIGKLVNRIVLLTSRDIYDIALCF 207  
QY 137 EKNVKFYKCGFSN 150  
DB 208 EDERPFFKACGFGD 221

## RESULT 8

US-10-471-571A-5116  
; Sequence 5116, Application US/10471571A  
; Publication No. US20060115490A1  
; GENERAL INFORMATION:  
; APPLICANT: CHIRON SpA  
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS  
; FILE REFERENCE: P026927W0  
; CURRENT APPLICATION NUMBER: US/10/471,571A  
; CURRENT FILING DATE: 2003-09-12  
; PRIOR APPLICATION NUMBER: GB-0107661.1  
; PRIOR FILING DATE: 2001-03-27  
; NUMBER OF SEQ ID NOS: 5642  
; SOFTWARE: Seqwin99, version 1.03  
; SEQ ID NO 5116  
; LENGTH: 149  
; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (1)..(149)  
; OTHER INFORMATION: conserved hypothetical  
US-10-471-571A-5116

Query Match 10.6%; Score 88.5; DB 6; Length 149;  
Best Local Similarity 24.8%; Pred. No. 0.085;

Matches 38; Conservative 22; Mismatches 50; Indels 43; Gaps 6;  
QY 3 LPDGFYIRRR---MEEGDLEQVETLKVLTVTGTTIPESFSLIKYWNENATVWNNEDKKI 59  
DB 16 LEDCFYIRKKYFVEEQGVPESE-----IDEVESI-HLIGYDNGQPV----- 58  
QY 60 MQYNPMVIVDKRTETVATGNIIRKIIHELGLCHIEDIAVNSKYQGGLKLLIDQL 119  
DB 59 -----ATARI---RPINETTV--KIERVAVMKSHRGQGMRLMQAV 95  
QY 120 VTIGFDYGCYKILDCDEKNVKFYKCGFSNAG 152  
DB 96 ESLAKDEGFYVATMNAOCHAIPFYESLNFMRG 128

## RESULT 9

US-10-370-959-55  
; Sequence 55, Application US/10370959  
; Publication No. US20060088907A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc  
; APPLICANT: Meyers, Rachel E.  
; APPLICANT: Williamson, Mark W.  
; APPLICANT: Kapeller-Libermann, Rosana  
; APPLICANT: MacBeth, Kyle J.  
; APPLICANT: Hunter, John J.  
; APPLICANT: Rudolph-Owen, Laura A.  
; APPLICANT: Bandaru, Rajasekhar  
; APPLICANT: Tsai, Fong-Ying  
; TITLE OF INVENTION: NOVEL 13237, 18480, 2245, 16228, 7677,  
; TITLE OF INVENTION: 26320, 46619, 33166, 16836, 46867, 21617, 55562, 39228,  
; TITLE OF INVENTION: 62088, 46745, 23155, 21657, 42755, 23229, 22325, 46863 AND  
; TITLE OF INVENTION: 32252 MOLECULES AND USES THEREFOR  
; FILE REFERENCE: MPI03-017OMNIM  
; CURRENT APPLICATION NUMBER: US/10/370,959  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: US 09/910,150  
; PRIOR FILING DATE: 2001-07-18  
; PRIOR APPLICATION NUMBER: US 60/219,028  
; PRIOR FILING DATE: 2000-07-18  
; PRIOR APPLICATION NUMBER: US 10/251,507  
; PRIOR FILING DATE: 2002-09-20  
; PRIOR APPLICATION NUMBER: US 09/715,479  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: US 60/218,053  
; PRIOR FILING DATE: 2000-07-13  
; PRIOR APPLICATION NUMBER: US 09/644,929  
; PRIOR FILING DATE: 2000-08-23  
; PRIOR APPLICATION NUMBER: US 60/212,439  
; PRIOR FILING DATE: 2000-06-16  
; PRIOR APPLICATION NUMBER: US 09/892,870  
; PRIOR FILING DATE: 2001-06-26  
; PRIOR APPLICATION NUMBER: US 60/214,174  
; PRIOR FILING DATE: 2000-06-26  
; PRIOR APPLICATION NUMBER: US 09/775,117  
; PRIOR FILING DATE: 2001-02-01  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 156  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 55  
; LENGTH: 139  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-10-370-959-55

Query Match 10.0%; Score 83.5; DB 6; Length 139;  
Best Local Similarity 31.6%; Pred. No. 0.24;  
Matches 18; Conservative 15; Mismatches 21; Indels 3; Gaps 1;

QY 95 GHIEDIAVNSKYQGGLKLLIDQLVTIGFDYGCYKILDCDEKN---VKFYKCGF 148  
DB 59 GYIAMLAVDSKYRRNGIGTNLVKKAIVAMVEGDCDEVVLEITNKSALKLYENLGF 115



```
Query Match          9.5%; Score 79.5; DB 6; Length 492;
Best Local Similarity 27.5%; Pred. No. 2.8; Mismatches 15; Gaps 3;
Matches 30; Conservative 15; Indels 9; Gaps 3;

QY 48 ATVMNDNEDKKIMQYNPMVIVDKR-TETVAATGNIHIERKIIHELGL-----CGH-IE 98
Db 23 APVMNDNEALTVGPRGPILLDDYHLIEKVAHFARERIPERVVHARGASAKGFFECTHDVT 82

QY 99 DIANSKYQGGGLKLLIDQLVITGFDYGCYKIIILDCDEKNVKFYKCG 147
Db 83 DITCADFLRSPCAQTPVIVRFSTVIHERGSPETIRDRGFAVKFYTREG 131

RESULT 14
US-10-471-571A-4442
; Sequence 4442, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927WO
; CURRENT APPLICATION NUMBER: US/10/471,571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 4442
; LENGTH: 133
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(133)
; OTHER INFORMATION: hypothetical protein
US-10-471-571A-4442

Query Match          9.5%; Score 79; DB 6; Length 133;
Best Local Similarity 31.3%; Pred. No. 0.62; Mismatches 11; Gaps 2;
Matches 26; Conservative 11; Indels 10; Gaps 2;

QY 72 TETVAATGNIHIERKIIHELGLCGHIEDIANSKYQGGGLKLLIDQL-----VTIGFD 125
Db 41 TVTLYDKDRLLGMRGVIGDGTVOIVDIANVLKSIQGGYGSLSLIMEHIMQIKGVAVEST 100

QY 126 YGCYKIIILDCDEKNVKFYKCGF 148
Db 101 Y----VSLIADYPADKLYTKGF 119

RESULT 15
US-10-953-349-10209
; Sequence 10209, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 10209
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-10209

Query Match          9.2%; Score 77; DB 6; Length 435;
Best Local Similarity 20.4%; Pred. No. 4.2; Mismatches 24; Indels 26; Gaps 4;
Matches 30; Conservative 24; Mismatches 67; Indels 26; Gaps 4;
```

```
QY 10 RMEEGLDQVETLKVLTITVTTPESFSKLIKYWNEATVMNDNEDKKIMQYNPMVIVD 69
Db 275 REAKVEDLAGIRQIIRKLEESGALVRRTDELLR-----ALDSFVVVE 317

QY 70 KRTETVAATGNI-IIERKIIHELGLCGHIEDIANSKYQGGGLKLLIDQLVITGFDYGC 128
Db 318 REGHIIACAALFPFFEEK-----CGEVAIAIVASDCRGQGGQKLLIDYIEKKASALGL 370

QY 129 YKIIILDCDEKNVKFYKCGFSNAGVEM 155
Db 371 EMLFL-LTTRTADWFRVRGFOECPIEM 396
```

Search completed: June 14, 2006, 15:34:54  
Job time : 6.38281 secs

**THIS PAGE IS BLANK**



Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	835	100.0	159	2	US-09-538-092-266	Sequence 266, Appl
2	328	39.3	155	2	US-09-248-796A-14732	Sequence 14732, A
3	209	25.0	184	2	US-09-795-926-46	Sequence 46, Appl
4	209	25.0	184	2	US-10-364-774-46	Sequence 46, Appl
5	194	23.2	114	2	US-09-270-767-32856	Sequence 32856, A
6	194	23.2	114	2	US-09-270-767-48073	Sequence 48073, A
7	131.5	15.7	153	2	US-09-270-767-31849	Sequence 31849, A
8	131.5	15.7	153	2	US-09-270-767-47066	Sequence 47066, A
9	117.5	14.1	242	2	US-09-253-991A-28218	Sequence 28218, A
10	110	13.2	179	2	US-09-603-208A-20	Sequence 20, Appl
11	106	12.7	149	2	US-09-107-532A-5296	Sequence 5296, Ap
12	104	12.5	307	2	US-09-543-681A-6849	Sequence 6849, Ap
13	101	12.1	170	1	US-08-742-009-1	Sequence 1, Appli
14	101	12.1	170	2	US-09-172-110-1	Sequence 1, Appli
15	98.5	11.8	86	2	US-09-134-000C-6710	Sequence 6710, Ap
16	98	11.7	23	2	US-09-927-734C-5	Sequence 5, Appli
17	93	11.1	126	2	US-09-902-540-15328	Sequence 15328, A
18	91.5	11.0	205	2	US-09-107-532A-5462	Sequence 5462, Ap
19	91.5	11.0	227	2	US-09-328-352-4162	Sequence 4162, Ap
20	91.5	11.0	556	1	US-08-503-377-1	Sequence 1, Appli
21	91.5	11.0	556	2	US-08-798-269-1	Sequence 1, Appli
22	91.5	11.0	556	2	US-09-055-210-1	Sequence 1, Appli
23	91.5	11.0	556	2	US-09-298-924-8	Sequence 8, Appli
24	91	10.9	148	2	US-09-823-414-17	Sequence 17, Appl
25	90	10.8	145	2	US-09-107-532A-7085	Sequence 7085, Ap
26	89.5	10.7	204	2	US-09-489-039A-8937	Sequence 8937, Ap

Query Match 25.0%; Score 209; DB 2; Length 184;  
Best Local Similarity 32.9%; Pred. No. 3.3e-17;  
Matches 48; Conservative 25; Mismatches 61; Indels 12; Gaps 2;

```

; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 32856
; LENGTH: 114
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-32856

Query Match      23.2%; Score 194; DB 2; Length 114;
Best Local Similarity 40.2%; Pred. No. 1.1e-15;
Matches 35; Conservative 18; Mismatches 34; Indels 0; Gaps 0;

QY 62 YNPMVIVDKRTETVAATGNIIRKIIHELGLCGHIEDIANVSKYQGGLGKLLIDQLVT 121
DB 17 YFTVIEDTRKNEIIGAASLVIERKFHNRAVRGRLEDVVNDTYRGKQGLKLIIVTVSL 76

QY 122 IGFYGCYKIILDCDKNVKFKYKCGF 148
DB 77 LAELGCGYKMSLDCKDKLIKFKYESLGY 103

RESULT 6
US-09-270-767-48073
; Sequence 48073, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 48073
; LENGTH: 114
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-48073

Query Match      23.2%; Score 194; DB 2; Length 114;
Best Local Similarity 40.2%; Pred. No. 1.1e-15;
Matches 35; Conservative 18; Mismatches 34; Indels 0; Gaps 0;

QY 62 YNPMVIVDKRTETVAATGNIIRKIIHELGLCGHIEDIANVSKYQGGLGKLLIDQLVT 121
DB 17 YFTVIEDTRKNEIIGAASLVIERKFHNRAVRGRLEDVVNDTYRGKQGLKLIIVTVSL 76

QY 122 IGFYGCYKIILDCDKNVKFKYKCGF 148
DB 77 LAELGCGYKMSLDCKDKLIKFKYESLGY 103

RESULT 7
US-09-270-767-31849
; Sequence 31849, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 31849
; LENGTH: 153
; TYPE: PRT

```

```

; ORGANISM: Drosophila melanogaster
US-09-270-767-31849

Query Match      15.7%; Score 131.5; DB 2; Length 153;
Best Local Similarity 29.6%; Pred. No. 7.5e-08;
Matches 34; Conservative 24; Mismatches 42; Indels 15; Gaps 4;

QY 9 IRRMEGDLQEQ-VTETLKVLTITVTGTTPESEKLIKVMNEATVMNDEKKNQYNPMVI 67
DB 51 VRPLKDTDYDRGFLQLLSQLTHVGNVNRTPF--LTRF-----SQMKASGDYFTVTI 99

QY 68 VDKRTETVAATGNIIRKIIHELGLCGHIEDIANVSKYQGGLGKLLIDQLVTI 122
DB 100 EDRKNEIIGAASLVIERKFHNCAVRGRLEDVVNDTYRGKQGLKLIIV---VTV 151

RESULT 8
US-09-270-767-47066
; Sequence 47066, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 47066
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-47066

Query Match      15.7%; Score 131.5; DB 2; Length 153;
Best Local Similarity 29.6%; Pred. No. 7.5e-08;
Matches 34; Conservative 24; Mismatches 42; Indels 15; Gaps 4;

QY 9 IRRMEGDLQEQ-VTETLKVLTITVTGTTPESEKLIKVMNEATVMNDEKKNQYNPMVI 67
DB 51 VRPLKDTDYDRGFLQLLSQLTHVGNVNRTPF--LTRF-----SQMKASGDYFTVTI 99

QY 68 VDKRTETVAATGNIIRKIIHELGLCGHIEDIANVSKYQGGLGKLLIDQLVTI 122
DB 100 EDRKNEIIGAASLVIERKFHNCAVRGRLEDVVNDTYRGKQGLKLIIV---VTV 151

RESULT 9
US-09-252-991A-28218
; Sequence 28218, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28218
; LENGTH: 242
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28218

Query Match      14.1%; Score 117.5; DB 2; Length 242;
Best Local Similarity 28.1%; Pred. No. 7.4e-06;
Matches 27; Conservative 22; Mismatches 44; Indels 3; Gaps 1;

```

```
Qy 65 MVIDKRTETVAATGNIIIEKIIHELGLCGHIEDIAVNSKYQGQGLKLLIDQLVTIGF 124
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 144 VTLVAERGERLLGCTCLHLIEHLAHDFAARSAILDMVDRHARGQGVGRELIIGRAVERAR 203
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 125 DYGCYKILDCD---EKNVKPYEKCGRSAGNEMQI 157
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 204 SWGCYKLLALSSHQDRETAQREYAALGFTSHGVSLAL 239
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 10
US-09-603-208A-20
; Sequence 20, Application US/09603208A
; Patent No. 6822084
; GENERAL INFORMATION:
; APPLICANT: Pompeius, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; APPLICANT: Lee, Heung-Shick
; APPLICANT: Kim, Hyung-Joon
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING STRESS,
; FILE REFERENCE: RESISTANCE AND TOLERANCE PROTEINS
; CURRENT APPLICATION NUMBER: US/09/603,208A
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142692
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: 60/151214
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19930429.7
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931413.6
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931457.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931541.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932209.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932230.9
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932914.1
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19940764.9
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19941382.7
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 306
; SEQ ID NO 20
; LENGTH: 179
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-603-208A-20

Query Match 13.2%; Score 110; DB 2; Length 179;
Best Local Similarity 27.2%; Pred. No. 4e-05;
Matches 44; Conservative 23; Mismatches 69; Indels 26; Gaps 7;

Qy 7 FYIRMESGDLQVETELKVLTVTTGTTIPESFKLIKYNEATVWNNEDKKIMQYNPMV 66
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 6 FTIRPIREGDPQVRDIYELGLETCHAT-----YETSGPTWDQFSQSKIND-TVMV 55
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 67 IVDKRTET-----VAATGNIIIEKIIHELGLCGHIED-IAVNSKYQGQGLKLLIDQLV 120
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 56 AVENNDPDPFILGWSAAP--ISSRQVFH-----GVVEDSIYTHPQOGGRGIGGALLDALI 108
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 121 TIGFDYGCYKJ---ILDCDEKNVKPYEKCGRSAGNEMQIR 159
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 109 TYCSNGIWSHSHWIFPENLGSAKLHESKGFVKVGTMQMAR 150
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
RESULT 11
US-09-107-532A-5296
; Sequence 5296, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 5296:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 149 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...149
; SEQUENCE DESCRIPTION: SEQ ID NO: 5296:
US-09-107-532A-5296

Query Match 12.7%; Score 106; DB 2; Length 149;
Best Local Similarity 36.1%; Pred. No. 9.4e-05;
Matches 26; Conservative 14; Mismatches 30; Indels 2; Gaps 1;

Qy 79 GNIIEKIIHELGLCGHIEDIAVNSKYQGQGLKLLIDQLVTIGFDYGCYKJ--ILDCD 136
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 52 GKIIIGVRWITDFATFIQDILIHPRYQRQGIKGVLLNEALEKITSYGPVQIELLTD 111
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 137 EKNVKPYEKCGRF 148
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 112 EKTKKFYESVGF 123
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 12
US-09-543-681A-6849
; Sequence 6849, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILI
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
```

```

US-08-742-009-1
Query Match      12.1%; Score 101; DB 1; Length 170;
Best Local Similarity 23.1%; Pred. No. 0.00046;
Matches 37; Conservative 26; Mismatches 55; Indels 42; Gaps 6

Qy 9 IRMEEGDLEQVETLKVLTVTGTTIPESFSKLIKYNEATVWNDNEDKKIMQYNPMVIV 68
   |||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db 6 IREAKEGDGCDGLRLIRLEAEF-----EKLSDQVKISEEARADFGDGNPF--YHCLV-- 56
   |||::||::||::||::||::||::||::||::||::||::||::||::||::||::||

Qy 69 DKRTETVAATGNIILIERKIHELGLCG-----HTEDIAVNSKYQGQ 110
   |||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db 57 ---AELPAPGKL-----LGPCVVGGIYYFYSTWKGRITYLEDIYVMFEYRQQ 104
   |||::||::||::||::||::||::||::||::||::||::||::||::||::||::||

Qy 111 LGKLLIDQLVITGFYDGYCYKI---ILDCDEKNVKFYEKCG 147
   |||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db 105 IGSKIHKKVAEVALDKGCOSFRFLAVLDWNORAMDLYKALG 144
   |||::||::||::||::||::||::||::||::||::||::||::||::||::||::||

```

RESULT 14  
US-09-172-110-1  
; Sequence 1, Application US/09172110  
; Patent No. 6017529  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; TITLE OF INVENTION: A NOVEL HUMAN SPERMIDINE/  
; TITLE OF INVENTION: SPERMINE N1-ACETYLTRANSFERASE  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
;

```

ADDRESS: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: US
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/172,110
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/742,009
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0146 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 170 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY: Consensus
CLONE:
US-09-172-110-1
Query Match 12.1%; Score 101; DB 2; Length 170;
Best Local Similarity 23.1%; Pred. No. 0.00046;
Matches 37; Conservative 26; Mismatches 55; Indels 42; Gaps 6
Qy 9 IRMEEGDLQGVTTETLKVLTVTGTTTPESFSKLIKYWNEATVWNNDKKIKMQNPMVIV 68
Db 6 IRRAKEDGCDGDIURLIRELAEF-----EKLSDOVKISEEALRADGFGDNPF--YHCLV-- 56

```

```
Qy 69 DKRTETVAATGNIIIRKIIHELGLCG-----HIEDIAVNSKYQGQG 110
Db 57 ---AEILPAPGKL-----LQPCVVGYGIVYFYIYTWKGRITYLEDIYVMEYRQG 104
Qy 111 LGKLLIDQLVTIGFDYGCYKI---ILDCDEKNVYFYKCG 147
Db 105 IGSKIIRKVAEVALDKGCSQRLAVLDWNRAMDLYKALG 144
```

```
RESULT 15
US-09-134-000C-6710
; Sequence 6710, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6710
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-6710
```

```
Query Match 11.8%; Score 98.5; DB 2; Length 86;
Best Local Similarity 33.3%; Pred. No. 0.00036;
Matches 23; Conservative 15; Mismatches 20; Indels 11; Gaps 3;
Qy 96 HIEDIAVNSKYQGGLKLLIDQLVTIGFDYGCYK---IILDCDEKNV---KPYEKCGF 148
Db 4 YIEDIAVCKDPRGGIGSALIN---ISIEWAKHNLHGLMLETQDNNLIACKFYHNCGF 59
Qy 149 SNAGVEMQI 157
Db 60 KIGSVDTWML 68
```

Search completed: June 14, 2006, 15:29:52  
Job time : 16.9062 secs



[illegible][illegible]



A:Experimental source: clone T23G11  
C:Genetics:  
A:Gene: CESP-T23G11.2  
A:Map position: 1  
A:Introns: 74/2; 128/3; 250/2

Query Match 23.0%; Score 192; DB 2; Length 347;  
Best Local Similarity 32.2%; Pred. No. 5.9e-10;  
Matches 48; Conservative 25; Mismatches 64; Indels 12; Gaps 2;

QY 1 MSLPDGYFIRMEEGDLEQVETLTKVLTGTTTPESFSKLKYWNEATVWMDNEDKKIM 60  
DB 16 ISVPTGLRLRLRNDDGYL-ELLKQLTSVGFNQLVFRKQF-----DAMKAK 63  
QY 61 QYNPMVVDKRTETVAATGNIIERKIIHELGLCGHIEDIANVSKYQOGGLKLLIDQLV 120  
DB 64 SYIVVLEHIESKGIAGATLLEFKYIIEAGTRGRVEDVVDKMRGKGLGALLNEVLV 123

QY 121 TIGFDYGCYKIILDCDEKNVFKYKCGFS 149  
DB 124 EMAKTIGVYKLSLECKTELIPFNKFGYS 152

RESULT 6  
AC1713  
Hypothetical protein homology lin2246 [imported] - Listeria innocua (strain Clip11262)  
C:Species: Listeria innocua  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 31-Dec-2004  
C:Accession: AC1713  
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A:Authors: Kreft, J.; Kuhn, M.; Kunet, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Makok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.  
Article: Comparative genomics of Listeria species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
C:Accession: AC1713  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-157 <GLA>  
A:Cross-references: UNIPROT:Q929M8; UNIPARC:UPI00000CC7D6; GB:AL592022; PIDN:CAC97475.1;  
A:Experimental source: strain Clip11262  
C:Genetics:  
A:Gene: lin2246  
C:Superfamily: ribosomal protein serine N-acetyltransferase

Query Match 15.0%; Score 125.5; DB 2; Length 157;  
Best Local Similarity 27.7%; Pred. No. 0.00028;  
Matches 44; Conservative 27; Mismatches 47; Indels 41; Gaps 6;

QY 9 IRRMEEGDLEQVETLTKVLTGTTTPESFSKLKYWNEATVWMDNEDKKIMQYNP--MV 66  
DB 3 IRLSKRDSASMIIEHLVWTPGT-TP-----GDVHFDSEAEFLKKNPPGSKI 49  
QY 67 IVDKRTETVAATGNIIERKIIHELGLCGHIE-----DIANVSKYQOGGLKLLI 116  
DB 50 VVEK-----DEKIIGILYKSPILPSNKHVVEIDIAVHPDYQREGIGQLLM 96

QY 117 DQVTVIGFDYCYKI---ILDCDEKNVFKYKCGFSNAG 152  
DB 97 DKMEVAREKGFIKISRLVLSINQAKIRFYKNGFKQEG 135

RESULT 7  
C71136  
Hypothetical protein PH0855 - Pyrococcus horikoshii  
C:Species: Pyrococcus horikoshii  
C:Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 09-Jul-2004  
C:Accession: C71136  
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin, M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi, DNA Res. 5, 55-76, 1998

A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic  
A:Reference number: A71000; MUID:98344137; PMID:9679194  
A:Accession: C71136  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-312 <KAW>  
A:Cross-references: UNIPROT:O58585; UNIPARC:UPI0000062F12; GB:AP000003; NID:G3236130; P  
A:Experimental source: strain OT3  
A:Note: this accession replaces an interim accession for a sequence replaced by GenBank  
C:Genetics:  
A:Gene: PH0855

Query Match 14.7%; Score 123; DB 2; Length 312;  
Best Local Similarity 27.9%; Pred. No. 0.0011;  
Matches 38; Conservative 24; Mismatches 50; Indels 24; Gaps 4;

QY 37 SPSKLKY-----WNEATVWMDNEDKKIMQYNPMVVDKRTETVAAT 78  
DB 23 SVSKVVKYINGKEVEAKYDDLTVARWSHGPGFMSLETCAINTN-LLINDQYPLVAELN 81  
QY 79 GNIITIERKI-IHE-----LGLCGHIEDIANVSKYQOGGLKLLIDQLVTFIDFYGCYKIL 133  
DB 82 GKIVGELELYIGEESKLLGCGYIDVLEVHKDYRKGVGKALVKNKAVEIAKEHECDTVAV 141

QY 134 DCDEKNVFKYKCGFS 149  
DB 142 WPKVKAAGFYRCKGIS 157

RESULT 8  
B83223  
Probable acetyltransferase PA3368 [imported] - Pseudomonas aeruginosa (strain PAO1)  
C:Species: Pseudomonas aeruginosa  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004  
C:Accession: B83223  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warriner, P.; Hickey, M.J.; B  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lin  
; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic path  
A:Reference number: A82950; MUID:20437337; PMID:10984043  
A:Accession: B83223  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-221 <STO>  
A:Cross-references: UNIPROT:Q9HYN3; UNIPARC:UPI00000C59D5; GB:AE004759; GB:AE004091; NI  
A:Experimental source: strain PAO1  
C:Genetics:  
A:Gene: PA3368

Query Match 14.2%; Score 118.5; DB 2; Length 221;  
Best Local Similarity 24.1%; Pred. No. 0.0018;  
Matches 40; Conservative 30; Mismatches 67; Indels 29; Gaps 5;

QY 3 LPD----GFYIRMEEGDLEQVETLTKVLTGTTTPESFSKLKYWNEATVWMDNEDKK 58  
DB 71 IPDSPFASFDVRIASEHDAPRLSLLQL---GSADPRDPALL-----A 112

QY 59 IMQYNP-----MVVDKRTETVAATGNIIERKIIHELGLCGHIEDIANVSKYQOGGLK 114  
DB 113 IQLQRPGRDVTVAERGERLLGTCTLHLEHLAHDFAISAILEDMMVVDVDRHARGGVGRE 172

QY 115 LIDQLVTFIDFYGCYKIILDCD---EKNVFKYKCGFSNAGVEMOI 157  
DB 173 LIGRAVERARSWGCGYKALSSHQDRETAQRFYAALGFTSHGVSLAL 218

RESULT 9  
AE1342  
Hypothetical protein lmo2141 [imported] - Listeria monocytogenes (strain EGD-e)  
C:Species: Listeria monocytogenes  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 31-Dec-2004  
C:Accession: AE1342

R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker  
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Duseurget, O.; Entian, K.D.; Fiehi, H.  
Science 294, 849-852, 2001

A:Authors: Kretz, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma  
ok, C.; Schluerer, T.; Smoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,  
A:Title: Comparative genomics of *Listeria* species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AE1342

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-157 <GUA>

A:Cross-references: UNIPROT:Q8Y5C4; UNIPARC:UPI0000055747; GB:NC\_003210; PIDN:CAD00219.1

A:Experimental source: strain EGD-e

C:Genetics:

C:Superfamily: ribosomal protein serine N-acetyltransferase

Query Match 14.0%; Score 116.5; DB 2; Length 157;

Best Local Similarity 26.0%; Pred. No. 0.0018; Mismatches 54; Indels 29; Gaps 6;

Matches 40; Conservative 31; Mismatches 54; Indels 29; Gaps 6;

QY 9 IRRMEGLQVTEFLKVLTTVGTITPESFSLIKYNEATVWMDNEDKKIMQYNP-----64

DB 3 IRLSKREDAASMLEHLVWTPGT-TP-----GDIHFDSEAEFLKSPGSKI 49

QY 65 MVIVDKRTETVAATGNII--IERKIIHELGLGHIEDIAVNSKYQGGLKLLIDOLVTI 122

DB 50 VVVVTDKVVGILGYKSPILASNKHVAI-----DIAVHPDYQORAGIGQLLMDKMKVEV 102

QY 123 GFDYCYKI--ILLDCENKVKFYKCGFSNAGV 153

DB 103 AREGYIKIALRVLSINQKAIKRFYKNGFKQEG 136

RESULT 10

G86699

hypothetical protein yfiL [imported] - Lactococcus lactis subsp. lactis (strain IL1403)

C:Species: Lactococcus lactis subsp. lactis

C>Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 09-Jul-2004

C:Accession: G86699

R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissensbach, J.; Ehrli

Genome Res. 11, 731-753, 2001

A:Title: The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* s

A:Reference number: AB6625; MUID:21235386; PMID:11337471

A:Accession: G86699

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-154 <STO>

A:Cross-references: UNIPROT:Q9CHW9; UNIPARC:UPI00000687C; GB:AE005176; PID:g12723494; H

A:Experimental source: strain IL1403

C:Genetics:

A:Gene: yfiL

Query Match 13.5%; Score 113; DB 2; Length 154;

Best Local Similarity 37.3%; Pred. No. 0.0038;

Matches 31; Conservative 17; Mismatches 25; Indels 10; Gaps 4;

QY 70 KRTETVAATGNIIERKIIHELGLGHIEDIAVNSKYQGGLKLLIDOLVT-IGFDYGC 128

DB 44 KNGKTVAC-AVVIKKQKTFE-----IENFAVETSGQFGQGMKQLITYLKENLAA 96

QY 129 YKILLDCDE--KNVKEVEKCGFS 149

DB 97 DELILGTDVSNNAVFYKCGFT 119

RESULT 11

E90074

conserved hypothetical protein SA2454 [imported] - Staphylococcus aureus (strain N315)

C:Species: Staphylococcus aureus

C>Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004

C:Accession: E90074

R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc  
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;  
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.

Lancet 357, 1225-1240, 2001

A:Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.

A:Reference number: AB9758; MUID:21311952; PMID:11418146

A:Accession: E90074

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-166 <KUR>

A:Cross-references: UNIPROT:Q9QX7; UNIPARC:UPI000000CA963; GB:BA000018; PID:g13702619; P

A:Experimental source: strain N315

C:Genetics:

A:Gene: SA2454

Query Match 13.5%; Score 113; DB 2; Length 166;

Best Local Similarity 22.3%; Pred. No. 0.0041;

Matches 37; Conservative 30; Mismatches 53; Indels 46; Gaps 6;

QY 9 IRRMEGLQVTEFLKVLTTVG-----TIPESFSLIKYNEATVWMDNEDKKIMQYN 63

DB 3 LRKVMQDLDOI-----ALENIGFSPPEAATREALKRIEQIET-----43

QY 64 PMVIVDKRTETVA-ATGNIIERKIIHELGLGHIEDIAVNS-----KXQ 107

DB 44 -FIVAENKNDVIGVINGPVIKERYISDDL-----FKVNSINSEGGYISVLGLVVPVNYQ 97

QY 108 GQGLGKLLIDOLVTIGFDYGCYKILLDCENKVKFYKCGFSNAGV 153

DB 98 GQGIAGRLNFTETLAKNHRHGVTLTCTRESLSIFYEKGYRNEGV 143

RESULT 12

C97300

probable acetyltransferase [imported] - Clostridium acetobutylicum

C:Species: Clostridium acetobutylicum

C>Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 09-Jul-2004

C:Accession: C97300

R;Nolling, J.; Berton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,

J. Bacteriol. 183, 4823-4838, 2001

A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium C

A:Reference number: A96900; MUID:21359325; PMID:21359325

A:Accession: C97300

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-140 <KUR>

A:Cross-references: UNIPROT:Q97B60; UNIPARC:UPI000000CA784; GB:AE001437; PIDN:AAK81190.1,

A:Experimental source: Clostridium acetobutylicum ATCC824

C:Genetics:

A:Gene: CAC3256

Query Match 13.4%; Score 112; DB 2; Length 140;

Best Local Similarity 27.1%; Pred. No. 0.0041;

Matches 39; Conservative 25; Mismatches 52; Indels 28; Gaps 6;

QY 31 GTITPESFSLIKY--WNEATVWMDNEDKKIMQYNPMVIVDKRTETVAATGNIIERKII 88

DB 7 GSLTEPVNSLAESVGWGHPSIDQVEEALKNSLY-----TACVVDGNSPIAMGRIV 57

QY 89 HELGLGHIEDIAVNSKYQGGLKLL-----IDOLVTIGFDYGCYKILLDCDEKNVK 141

DB 58 GDLMSYFTKDVAVNPEYQNRIGKLLMDMLSFIESKTPPTGWK-TCIELLSAHGEG-- 114

QY 142 FYEKGCGF-----SNAGVEMOIR 158

DB 115 FYEKGFGQKRIKEKNGCGTLLIR 138

RESULT 13

E97249

probable acetyltransferase [imported] - Clostridium acetobutylicum

C:Species: Clostridium acetobutylicum

C: Superfamily: Escherichia coli ribosomal-protein-alanine N-acetyltransferase rimi  
A: Gene: CAC2840

```

QY 00 AINHEGDCGHIZIAVNSAKVQQGSGKRLADQVWILGDFGACRLIILDDGNNV-----144
Db 63 -IIDE---GHITIAVHPFRGFGIGNKILEELIKL-----CBKRNIPSMTL 105
QY 141 -----KFYEKCGFSNAGV 153
Db 106 EVRISNTIAONLYKKFGFGKEAGV 128

```

Search completed: June 14, 2006, 15:28:35  
Job time : 12.7656 secs

Qy	13	BEGDLEQVETFLKVLVTITTPESFSKLIIYWNBEATVMNDNEDKIIMYOYNPMIVDKRT	72
Dd	62	KEGDIAGV-----LAGYPGKIEPI-----DRAWNEIAKHGRHYEEPIFADKET	106
Qy	73	ETVAATGNIIIERKIIHBLGCGHIEDIAVNSKYQGGLGKLLIDQLVTIGFDYGCYKII	132
Dd	107	----FPGE-----WYLDSTVTEKKRGHGVTALLAKLITEIAADECKVKVG	148

Qy	133	LDDEKN-----VKFYKCGFSNAG	152
		:	
		:	
		:	
Dp	149	LNCDKGNPHAKRLYERLGFHVTG	171

QY	141	-----KFYEKCGFSNAGV	153
QY		: : :	
Db	106	EVISNTIAQNLYKKFGFKEAGV	128
Db		: : :	
Db	63	-IIDE---GHITNIAVHPEFRGMGINKLEBELIKL-----	CEKNIPSWTL 105
Db		: : :	
QY	86	KIHELGLCGHIEDIAVNSKYQGGLGKLIIDQLVTIGFYGCVKIILDCDEKNV-----	140
QY		: : :	
Db	5	TVHTLEKDDIDSIIETIENLCFPTPWTKESMEGELRNKFAKYVVIKNNLVVYGGGLW----	62
Db		: : :	

```

Query Match      12.7%; Score 106; DB 2; Length 185;
Best Local Similarity 25.2%; Pred. No. 0.021;
Matches 36; Conservative 20; Mismatches 51; Indels 36; Gaps 5;

Qy 13 EBGDLEOVTTKLVLTVGTITPESFKLIKYNEATVWVNDNEDKKIMQNPVIVDKRT 72
      |||: | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 62 KEGDIAGV----LAGYPGKIEPEI-----DRAWNEIAKKGIRVEEPIFADKET 106

Qy 73 ETVAATGNIIIERKIIHELGLCGHIEDIANSKYQGGGLKLLIDQLTVTFDYGCVKII 132
      :|: | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 107 ----FPGE-----WYDLSIVNEKYRGHGVGTALLAKTETAADGGEKVG 148

```

Qy	133	LDCEKN----	VKFYEKCGFSNAG	152
		: : :	: : :	
Db	149	LNCDKGNPHAKRLYERLGFHVITG		171

**THIS PAGE IS BLANK**

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 14, 2006, 15:15:09 ; Search time 79.707 Seconds  
(without alignments)  
1845.226 Million cell updates/sec

Title: US-10-612-779-30

Perfect score: 835

Sequence: 1 MSLPDGFYIRMEEGDLEQV.....VKFYKCGFSNAGVEMQIRK 159

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 7.2.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	835	100.0	159	1	GNAL YEAST
2	476.5	57.1	160	2	Q6FW86 CANGA
3	434.5	52.0	171	2	Q752R3 ASHGO
4	430.5	51.6	188	2	Q6CRN3 KLULA
5	347	41.6	150	2	Q6BX84 DEBHA
6	336	40.2	171	2	Q6C8F2 YARLI
7	333	39.9	149	2	Q5AHF9 CANAL
8	328	39.3	149	1	GNAL CANAL
9	314	37.6	170	2	Q2UGQ9 ASPOR
10	290	34.7	157	2	Q54WR8 DICDI
11	279.5	33.5	177	2	Q7SHA4 NEUCR
12	276	33.1	176	2	Q411G4 GIBZE
13	275	32.9	190	2	Q4WCU5 ASPFU
14	266	31.9	165	2	Q5SR34 CRYNE
15	266	31.9	165	2	Q5KF35 CRYNE
16	264	31.6	205	2	Q7QTX4 GIALA
17	264	31.6	205	2	Q86F18 GIALA
18	261.5	31.3	149	2	Q9LFU9 ARATH
19	249	29.8	177	2	Q4IE15 GIBZE
20	246.5	29.5	181	2	Q2KFU6 MAGGR
21	236	28.3	148	2	Q7ROY7 GIALA
22	235	28.1	159	2	Q5ONZ5 ENTMO
23	234.5	28.1	158	2	Q5IAE1 ENTHI
24	232	27.8	111	1	GNAL SCHPO
25	226	27.1	234	2	Q5ASM4 EMENT
26	222.5	26.6	333	2	Q4PIQ6 USTWA
27	216	25.9	203	2	Q8SRG4 ENCCU
28	215	25.7	182	2	Q86QP1 BRABE
29	214.5	25.7	165	1	GNAL CAEEL
30	209.5	25.1	1100	2	Q5KF71 CRYNE
31	209.5	25.1	1127	2	Q55QZ7 CRYNE

32	208.5	25.0	179	2	Q4SAJ2 TETNG
33	208	24.9	159	2	Q5ORB0_ENTHI
34	208	24.9	184	1	GNAL_HUMAN
35	208	24.9	184	1	GNAL_PONPY
36	207	24.8	215	2	Q8IMK5 DROME
37	207	24.8	219	1	GNAL DROME
38	206	24.7	148	2	Q4DGL9_TRYCR
39	206	24.7	148	2	Q4DXL6_TRYCR
40	205.5	24.6	148	1	GNAL_MIMIV
41	204	24.4	184	2	Q4VBJ4_BRARE
42	201.5	24.1	198	2	Q7Q212 ANOGA
43	201	24.1	184	1	GNAL_MOUSE
44	201	24.1	349	2	Q6OW31_CARBR
45	200	24.0	212	2	Q5Q126_AEDAE

ALIGNMENTS

RESULT 1	
GNAL_YEAST	STANDARD; PRT; 159 AA.
AC P43577;	
DT 01-NOV-1995,	integrated into UniProtKB/Swiss-Prot.
DT 07-MAR-2006,	entry version 52.
DE	Glucosamine 6-phosphate N-acetyltransferase (EC 2.3.1.4)
DE	(Phosphoglucosamine transacetylase) (Phosphoglucosamine acetylase).
GN	Name=GNAL; Synonyms=PAT1; OrderedLocusNames=YFL017C;
OS	Saccharomyces cerevisiae (Baker's Yeast).
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC	Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX	NCBI_TaxID=4933;
RN	[1]
RP	NUCLEOTIDE SEQUENCE [GENOMIC DNA], CHARACTERIZATION, AND MUTAGENESIS.
RX	MEDLINE=99085039; PubMed=9867860; DOI=10.1074/jbc.274.1.424;
RA	Mio T., Yamada-Okabe T., Arisawa M., Yamada-Okabe H.;
RT	"Saccharomyces cerevisiae GNAL, an essential gene encoding a novel acetyltransferase involved in UDP-N-acetylglucosamine synthesis.";
RL	J. Biol. Chem. 274:424-429(1999).
RN	[2]
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC	STRAIN=S288c / AB972;
RX	MEDLINE=95400292; PubMed=7670463;
RA	Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,
RA	Sasanuma S., Sasanuma M., Tsuchiya Y., Soeda E., Yokoyama K.,
RA	Yamazaki M., Tashiro H., Eki T.;
RT	"Analysis of the nucleotide sequence of chromosome VI from Saccharomyces cerevisiae.";
RL	Nat. Genet. 10:261-268(1995).
RN	[3]
RP	NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RC	STRAIN=S288c / AB972;
RA	Barrell B.G., Churcher C., Rajandream M.A.;
RL	Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
RN	[4]
RP	NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RC	STRAIN=S288c;
RA	Marsischky G., Rolfs A., Richardson A., Kane M., Baqui M., Taycher E.,
RA	Hu Y., Vannberg F., Weger J., Kramer J., Moreira D., Kelley F.,
RA	Zuo D., Raphael J., Hogle C., Jepson D., Williamson J., Camargo A.,
RA	Gonzaga L., Vasconcelos A.T., Simpson A., Kolodner R., Harlow E.,
RA	LaBaer J.;
RT	"Creation of the YFLEX clone resource: cloning of Saccharomyces cerevisiae ORFs in the Gateway recombinational cloning system.";
RL	Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
RN	[5]
RP	LEVEL OF PROTEIN EXPRESSION
RX	MEDLINE=22923965; PubMed=14562106; DOI=10.1038/nature02046;
RA	Ghaemmaghami S., Huh W.-K., Bower K., Howson R.W., Belle A.,
RA	Dephour N., O'Shea E.K., Weissman J.S.;
RT	"Global analysis of protein expression in yeast.";
RL	Nature 425:737-741(2003).



```
DB 121 TEIGFDAGCYKVILDCDEKNVAFYEKCGYKAGVEMQCR 159

RESULT 3
ID Q752R3 ASHGO PRELIMINARY; PRT; 171 AA.
AC Q752R3;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 16.
DE AFR510WP.
GN OrderedLocusNames=AFR510W;
OS Ashbya gossypii (Yeast) (Eremothecium gossypii).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Eremothecium.
OX NCBI_TaxID=33169;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=ATCC 10895;
RX PubMed=15001715; DOI=10.1126/science.1095781;
RA Dietrich F.S., Voegeli S., Brachat S., Lerch A., Gates K., Steiner S.,
RA Mohr C., Poehmann R., Luedi P., Choi S., Wing R.A., Flavler A.,
RA Gaffney T.D., Philippsen P.;
RT "The Ashbya gossypii genome as a tool for mapping the ancient
RT Saccharomyces cerevisiae genome.";
RL Science 304:304-307 (2004).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF016819; AAS53881.1; -; Genomic_DNA.
DR AGD; AFR510W; -;
DR GO; GO:0008080; F:N-acetyltransferase activity; IEA.
DR InterPro; IPR000182; GCN5acetyl_trans.
DR Pfam; PF00583; Acetyltransf_1; 1.
KW Complete proteome.
SQ SEQUENCE 171 AA; 18855 MW; 69F2E5310C2A186D CRC64;

Query Match 52.0%; Score 434.5; DB 2; Length 171;
Best Local Similarity 56.8%; Pred. No. 2.7e-31;
Matches 92; Conservative 16; Mismatches 47; Indels 7; Gaps 2;

QY 3 LPDGFYIRMEEGDLEQVETLTKVLTGVTIPESFKLIKYWNEATVWVNDNEDKK----58
DB 4 LPEGHVIRAEADYAGVETLTKVLTGVDGTFEREFAERIAVWKTVKVPVARGKRPVGM 63
QY 59 --IMQNPVIVDKRTETVAATGNIIRKIIHELGLCGHIEDIIVNSKYQGGLGLKLI 116
DB 64 GEILAYNPVITDE-AGRVVATGNIIRKIIHELGLCGHIEDIIVASDQQKRLGLLI 122
QY 117 DQVLTGFDYGYCKIILDCDEKNVAFYEKCGFSNAGVEMQIR 158
DB 123 NTLTEIRNAGCYKVIILDCDQPNADYFKKCGFSQAGLEMQHR 164

RESULT 4
Q6CRN3_KLULA PRELIMINARY; PRT; 188 AA.
AC Q6CRN3;
DT 16-AUG-2004, integrated into UniProtKB/TrEMBL.
DT 16-AUG-2004, sequence version 1.
DT 07-FEB-2006, entry version 11.
DE Kluyveromyces lactis strain NRRL Y-1140 chromosome D of strain NRRL Y-
DE 1140 of Kluyveromyces lactis.
GN OrderedLocusNames=KLIA0D07700g;
OS Kluyveromyces lactis (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
OX NCBI_TaxID=28985;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=CBS 2359 / IFO 1267 / NRRL Y-1140 / WM37;
RX PubMed=15229592; DOI=10.1038/nature02579;
```

```
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talla E.,
RA Coffard N., Frangoul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.-M., Beyne E., Bleykasten C.,
RA Boisarame A., Boyer J., Cattolico L., Confaniolero F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.-M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.-F., Straub M.-L., Suleau A.,
RA Swennen D., Tekia F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic Y., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.-L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44 (2004).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; CR382124; CAH00502.1; -; Genomic_DNA.
DR GO; GO:0008080; F:N-acetyltransferase activity; IEA.
DR InterPro; IPR000182; GCN5acetyl_trans.
DR Pfam; PF00583; Acetyltransf_1; 1.
KW Complete proteome.
SQ SEQUENCE 188 AA; 21116 MW; ED656D6A8C849071 CRC64;

Query Match 51.6%; Score 430.5; DB 2; Length 188;
Best Local Similarity 53.8%; Pred. No. 6.9e-31;
Matches 85; Conservative 21; Mismatches 45; Indels 7; Gaps 2;

QY 1 MSLPDGFYIRMEEGDLEQVETLTKVLTGVTIPESFKLIKYWNEATVWVNDNEDKKIM 60
DB 35 MSLPEGYTIRRTKDDFGSVTLKVTMGVDVQDFHSLDHDVSTIGD-----IP 88
QY 61 QYNPMVIVDKRTETVAATGNIIRKIIHELGLCGHIEDIIVNSKYQGGLGLKLI 120
DB 89 MYNSTVIVDTNGE-VVATGNVLVEKVIHECGLVGHIEDIIVAKDQGGKGLGLLIQYLY 147
QY 121 TGFYDYGCVKIIILDCDEKNVAFYEKCGFSNAGVEMQIR 158
DB 148 KLANEYGCYKVIILDCDESNVGFYKCGLKKAGVEMQIR 185

RESULT 5
Q6BX84_DEBHA PRELIMINARY; PRT; 150 AA.
AC Q6BX84;
DT 16-AUG-2004, integrated into UniProtKB/TrEMBL.
DT 16-AUG-2004, sequence version 1.
DT 07-FEB-2006, entry version 10.
DE Similar to sp|O93806 Candida albicans Glucosamine 6-phosphate N-
DE acetyltransferase.
GN OrderedLocusNames=DEHA0B05005g;
OS Debaryomyces hansenii (Yeast) (Torulaspora hansenii).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Debaryomyces.
OX NCBI_TaxID=4959;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=ATCC 36239 / CBS 767;
RX PubMed=15229592; DOI=10.1038/nature02579;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talla E.,
RA Coffard N., Frangoul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.-M., Beyne E., Bleykasten C.,
RA Boisarame A., Boyer J., Cattolico L., Confaniolero F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.-M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.-F., Straub M.-L., Suleau A.,
RA Swennen D., Tekia F., Wesolowski-Louvel M., Westhof E., Wirth B.,
```









[illegible]

Search completed: June 14, 2006, 15:27:38  
Job time : 82.707 secs

Result No.	Score	Query Match	Length	DB	ID	Description	
1	3089	100.0	609	7	US-11-245-473-22	Sequence 22, Appl	
2	3072	99.4	609	7	US-11-245-473-28	Sequence 28, Appl	
3	3072	99.4	609	7	US-11-245-473-31	Sequence 31, Appl	
4	3068	99.3	609	7	US-11-245-473-16	Sequence 16, Appl	
5	3061	99.1	609	7	US-11-245-473-25	Sequence 25, Appl	
6	3053	98.8	609	7	US-11-245-473-19	Sequence 19, Appl	
7	1145	37.1	601	6	US-10-471-571A-3268	Sequence 2268, Ap	
8	1104	35.7	681	6	US-10-480-021-6	Sequence 6, Appli	
9	1104	35.7	681	6	US-10-480-021-7	Sequence 7, Appli	
10	1098.5	35.6	682	6	US-10-480-021-8	Sequence 8, Appli	
11	1078.5	34.9	680	6	US-10-953-349-34174	Sequence 34174, A	
12	1071	34.7	697	6	US-10-449-902-38427	Sequence 38427, A	
13	801	25.9	491	6	US-10-953-349-34175	Sequence 34175, A	
14	800	25.9	488	6	US-10-953-349-34176	Sequence 34176, A	
15	582.5	18.9	464	6	US-10-488-015-17	Sequence 17, Appl	
16	551.5	17.9	286	6	US-10-449-902-36965	Sequence 36965, A	
17	252.5	8.2	541	6	US-10-449-902-54059	Sequence 54059, A	
18	241	7.8	501	6	US-10-471-571A-5078	Sequence 5078, Ap	
19	190.5	6.2	600	6	US-10-449-902-38704	Sequence 38704, A	
20	184.5	6.0	78	6	US-10-488-015-16	Sequence 16, Appl	
21	156.5	5.1	591	6	US-10-449-902-53151	Sequence 53151, A	
22	152.5	4.9	584	6	US-10-953-349-4929	Sequence 4929, Ap	
23	143	4.6	54	6	US-10-488-015-18	Sequence 18, Appl	
24	140	4.5	582	6	US-10-953-349-22931	Sequence 22931, A	
25	134	4.3	73	6	US-10-488-015-19	Sequence 19, Appl	

Db 241 YDAGDKGIYCHYMQKEIYEQPNAIKNTLTGRISHGQVDLSGLGNADLLSKVEHIQIIA 300  
Qy 301 CGTSNSGMVSRYPFESLAGIPCDVEIASBFRYRKSAVRRNSLMTLSQSGETADTLA 360  
Db 301 CGTSNSGMVSRYPFESLAGIPCDVEIASBFRYRKSAVRRNSLMTLSQSGETADTLA 360  
Qy 361 RLSKELGYLGLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQLTVLLMLVAKL 420  
Db 361 RLSKELGYLGLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQLTVLLMLVAKL 420  
Qy 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEDESDKHAFILSRGQOYPIA 480  
Db 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEDESDKHAFILSRGQOYPIA 480  
Qy 481 LEGALKLKEISYIHAEAYAAAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEVRAR 540  
Db 481 LEGALKLKEISYIHAEAYAAAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEVRAR 540  
Qy 541 GGQLYVFADODAGFVSSDNMHIIEMPHVEEVIAPFYTVPLQLLAYHVALIKGTVDQPR 600  
Db 541 GGQLYVFADODAGFVSSDNMHIIEMPHVEEVIAPFYTVPLQLLAYHVALIKGTVDQPR 600  
Qy 601 NLAKSVTVE 609  
Db 601 NLAKSVTVE 609

## RESULT 2

US-11-245-473-28  
; Sequence 28, Application US/11245473  
; Publication No. US20060094085A1  
; GENERAL INFORMATION:  
; APPLICANT: Berry, Alan  
; APPLICANT: Burlingame, Richard P.  
; TITLE OF INVENTION: PROCESS AND MATERIALS FOR PRODUCTION OF GLUCOSAMINE  
; FILE REFERENCE: 3161-18-C1  
; CURRENT APPLICATION NUMBER: US/11/245,473  
; CURRENT FILING DATE: 2005-10-05  
; PRIOR FILING DATE: 2001-12-17  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/115,475  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-15  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/035,494  
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-01-14  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 28  
; LENGTH: 609  
; TYPE: PRT  
; ORGANISM: Escherichia coli  
; US-11-245-473-28

Query Match 99.4%; Score 3072; DB 7; Length 609;  
Best Local Similarity 99.7%; Pred. No. 1.6e-219;  
Matches 607; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 MCGIVGAIAQRDVAEIILLEGRLRLEYRGYDSAGLAVVDTEGHMTLRRLRGKVQMLAQAAE 60  
Db 1 MCGIVGAIAQRDVAEIILLEGRLRLEYRGYDSAGLAVVDTEGHMTLRRLRGKVQMLAQAAE 60  
Qy 61 EHPHGGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHNEPLREELKARGYTFV 120  
Db 61 EHPHGGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHNEPLREELKARGYTFV 120  
Qy 121 SETDTEVIAHLVNWELKOGGTLREAVLRAIPOLRGAYGTVMDSRHPDTLLAARSGSPV 180  
Db 121 SETDTEVIAHLVNWELKOGGTLREAVLRAIPOLRGAYGTVMDSRHPDTLLAARSGSPV 180  
Qy 181 IGLGMGENFIASDQALLPVTRRRFTFLEBGDIAETRRSVNIFDKTGAEVKQDIESNLQ 240  
Db 181 IGLGMGENFIASDQALLPVTRRRFTFLEBGDIAETRRSVNIFDKTGAEVKQDIESNLQ 240

Qy 241 YDAGDKGIYCHYMQKEIYEQPNAIKNTLTGRISHGQVDLSGLGNADLLSKVEHIQIIA 300  
Db 241 YDAGDKGIYCHYMQKEIYEQPNAIKNTLTGRISHGQVDLSGLGNADLLSKVEHIQIIA 300  
Qy 301 CGTSNSGMVSRYPFESLAGIPCDVEIASBFRYRKSAVRRNSLMTLSQSGETADTLA 360  
Db 301 CGTSNSGMVSRYPFESLAGIPCDVEIASBFRYRKSAVRRNSLMTLSQSGETADTLA 360  
Qy 361 RLSKELGYLGLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQLTVLLMLVAKL 420  
Db 361 RLSKELGYLGLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQLTVLLMLVAKL 420  
Qy 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEDESDKHAFILSRGQOYPIA 480  
Db 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEDESDKHAFILSRGQOYPIA 480  
Qy 481 LEGALKLKEISYIHAEAYAAAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEVRAR 540  
Db 481 LEGALKLKEISYIHAEAYAAAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEVRAR 540  
Qy 541 GGQLYVFADODAGFVSSDNMHIIEMPHVEEVIAPFYTVPLQLLAYHVALIKGTVDQPR 600  
Db 541 GGQLYVFADODAGFVSSDNMHIIEMPHVEEVIAPFYTVPLQLLAYHVALIKGTVDQPR 600  
Qy 601 NLAKSVTVE 609  
Db 601 NLAKSVTVE 609

## RESULT 3

US-11-245-473-31  
; Sequence 31, Application US/11245473  
; Publication No. US20060094085A1  
; GENERAL INFORMATION:  
; APPLICANT: Berry, Alan  
; APPLICANT: Burlingame, Richard P.  
; TITLE OF INVENTION: PROCESS AND MATERIALS FOR PRODUCTION OF GLUCOSAMINE  
; FILE REFERENCE: 3161-18-C1  
; CURRENT APPLICATION NUMBER: US/11/245,473  
; CURRENT FILING DATE: 2005-10-05  
; PRIOR FILING DATE: 2001-12-17  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/115,475  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-15  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/035,494  
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-01-14  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 31  
; LENGTH: 609  
; TYPE: PRT  
; ORGANISM: Escherichia coli  
; US-11-245-473-31

Query Match 99.4%; Score 3072; DB 7; Length 609;  
Best Local Similarity 99.7%; Pred. No. 1.6e-219;  
Matches 607; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 MCGIVGAIAQRDVAEIILLEGRLRLEYRGYDSAGLAVVDTEGHMTLRRLRGKVQMLAQAAE 60  
Db 1 MCGIVGAIAQRDVAEIILLEGRLRLEYRGYDSAGLAVVDTEGHMTLRRLRGKVQMLAQAAE 60  
Qy 61 EHPHGGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHNEPLREELKARGYTFV 120  
Db 61 EHPHGGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHNEPLREELKARGYTFV 120  
Qy 121 SETDTEVIAHLVNWELKOGGTLREAVLRAIPOLRGAYGTVMDSRHPDTLLAARSGSPV 180  
Db 121 SETDTEVIAHLVNWELKOGGTLREAVLRAIPOLRGAYGTVMDSRHPDTLLAARSGSPV 180  
Qy 181 IGLGMGENFIASDQALLPVTRRRFTFLEBGDIAETRRSVNIFDKTGAEVKQDIESNLQ 240  
Db 181 IGLGMGENFIASDQALLPVTRRRFTFLEBGDIAETRRSVNIFDKTGAEVKQDIESNLQ 240



121	Db	121	SETDTEVIAHLVNWELKGGTLREAVURAI	PQLRGAYGTVMDSRHPDPTLLAARSGSPLV	180		
181	Qy	181	IGLGMGNFIASDQALLPVTRRFIFLEEGDIA	EITRRSVNI	FDKTCAEVKRQDIESNLQ	240	
181	Db	181	IGLGMGNFIASDQALLPVTRRFIFLEEGDIA	EITRRSVNI	FDKTCAEVKRQDIESNLQ	240	
241	Qy	241	YDAGDKGIYCHYMQKEIYEOPNAIKNTLGT	GRISHGQVDLS	SELGNADELLKSKEVHIQIIA	300	
241	Db	241	YDAGDKGIYRHYYMQKEIYEOPNAIKNTLGT	GRISHGQVDLS	SELGNADELLKSKEVHIQIIA	300	
301	Qy	301	CGTSYNSGMVSRWYFESLAGI	PCDVEIAS	EFYRKSAVRNSLMITLSQSGETADTLAGL	360	
301	Db	301	CGTSYNSGMVSRWYFESLAGI	PCDVEIAS	EFYRKSAVRNSLMITLSQSGETADTLAGL	360	
361	Qy	361	RLSKELGYLGSLAI	CNVPGSSLV	RESDLALMTNAGTEIGVASTKAF	TQTLVLLMLVAKL	420
361	Db	361	RLSKELGYLGSLAI	CNVPGSSLV	RESDLALMTNAGTEIGVASTKAF	TQTLVLLMLVAKL	420
421	Qy	421	SRLKGLDASIEHDI	VHGLQALPSRI	TEOMLSODKRIEALAE	DFSDKHALFLSRSGDQYPIA	480
421	Db	421	SRLKGLDASIEHDI	VHGLQALPSRI	TEOMLSODKRIEALAE	DFSDKHALFLSRSGDQYPIA	480
481	Qy	481	LEGALKLKEISYTHAEYAAGELKHGPAL	ADAMPVIVVAP	NNNELLEKLSKSNIEEVRAR	540	
481	Db	481	LEGALKLKEISYTHAEYAAGELKHGPAL	ADAMPVIVVAP	NNNELLEKLSKSNIEEVRAR	540	
541	Qy	541	GGLQYVFADQDAGFVSSDNMHI	IEMPHVEEVI	APIFTVPQLQALAHYHVALIKGTDDVQDR	600	
541	Db	541	GGLQYVFADQDAGFVSSDNMHI	IEMPHVEEVI	APIFTVPQLQALAHYHVALIKGTDDVQDR	600	
601	Qy	601	NLAKSVTVE	609			
601	Db	601	NLAKSVTVE	609			

## RESULT 6

```

US-11-245-473-19
; Sequence 19, Application US/11245473
; Publication No. US2006094085A1
; GENERAL INFORMATION:
; APPLICANT: Bexry, Alan
; APPLICANT: Burlingame, Richard P.
; APPLICANT: Mullis, James R.
; TITLE OF INVENTION: PROCESS AND MATERIALS FOR PRODUCTION OF GLUCOSAMINE
; FILE REFERENCE: 3161-18-C1
; CURRENT APPLICATION NUMBER: US/11/245,473
; CURRENT FILING DATE: 2005-10-05
; PRIOR APPLICATION NUMBER: US/10/024,460
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/115,475
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/035,494
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-01-14
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Escherichia coli
US-11-245-473-19

```

Qy	121	SETDTEVIAHLVNWELKQGQGTLEAVLRAIPQLRGAYGTVMDSRRHPTDTLLAARSGSPLV	180
Db	121	SETDTEVIAHLVNWELKQGQGTLEAVLRAIPQLRGAYGTVMDSRRHPTDTLLAARSGSPLV	180
Qy	181	IGLGMGNFTASDQALLPVTTRPFIPLFEEGDIAEITRRSVNIPDKTGAEVKRQDIESNLQ	240
Db	181	IGLGMGNFTASDQALLPVTTRPFIPLFEEGDIAEITRRSVNIPDKTGAEVKRQDIESNLQ	240
Qy	241	YDAGDKGIYCHYMQKEIYEOPNAIKNTLTGRI SHGQVDLSGLGPNADELLISKVEHIQILA	300
Db	241	YDAGDKGIYRHVMQKEIYEOPNAIKNTLTGRTSHGQVDLSGLGPNADELLISKVEHIQILA	300
Qy	301	CGTSYNGMSRVYWFESLAGIPCDVETASFPREKSAVRNSLMTITLSQSETADTLAGL	360
Db	301	CGTSYNGMSRVYWFESLAGIPCDVETASEFRYKRSAVRRNSLMTITLSQSETADTLAGL	360
Qy	361	RLSKELGYLGSGLAICNVPGSSLVRESDLMTNAGTEIGVASTKAFTTQLTVLLMLVAKL	420
Db	361	RLSKELGYLGSGLAICNVPGSSLVRESDLMTNAGTEIGVASTKAFTTQLTVLLMLVAKL	420
Qy	421	SRUKGLDASTEHDIVHGLQALPSRIEOMLSQDKIEALAEAFDSKHHALFLSRGDOYPIA	480
Db	421	SRUKGLDASTEHDIVHGLQALPSRIEOMLPQDKRIEALAEAFDSKHHALFLGRGDQYPIA	480
Qy	481	LEGALKLKEISYIHAEAYAAAGELKHGFLALIDADMPVIVVAPNNLEFLKKSNIIEEVRAR	540
Db	481	LEGALKLKEISYIHAEAYAAAGELKHGFLALIDADMPVIVVAPNNLEFLKKSNIIEEVRAR	540
Qy	541	GGQLYVFADODAGFVSSDNMHIITEMPVHEEVIAPFVTVPLQLLAYHVALLKGTVDQDPR	600
Db	541	GGQLYVFADODAGFVSSDNMHIITEMPVHEEVIAPFVTVPLQLLAYHVALLKGTVDQDPR	600
Qy	601	NLAKSVTVE 609	
Db	601	NLAKSVTVE 609	
RESULT 7			
US-10-471-571A-2268			
; Sequence 2268, Application US/10471571A			
; Publication No. US20060115490A1			
; GENERAL INFORMATION:			
; APPLICANT: CHIRON SPA			
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS			
; FILE REFERENCE: P026927WO			
; CURRENT APPLICATION NUMBER: US/10/471.571A			
; CURRENT FILING DATE: 2003-09-12			
; PRIOR APPLICATION NUMBER: GB-0107661.1			
; PRIOR FILING DATE: 2001-03-27			
; NUMBER OF SEQ ID NOS: 5642			
; SOFTWARE: SeqWin99, version 1.03			
; SEQ ID NO 2268			
; LENGTH: 601			
; TYPE: PRT			
; ORGANISM: Staphylococcus aureus			
US-10-471-571A-2268			

## RESULT 7

```

US-10-471-571A-2268
; Sequence 2268, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927WO
; CURRENT APPLICATION NUMBER: US/10/471,571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-01076661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: Seqwin99, version 1.03
; SEQ ID NO 2268
; LENGTH: 601
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-471-571A-2268

```









[illegible]

```

RESULT 14
US-10-953-349-34176
; Sequence 34176, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 34176
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
US-10-953-349-34176

```

Db	182	YNAALARPFPVEELTGTPVTHEVASDLLDRGPIYREDTAVFVSOGSETADTLLALDYAL	241
Qy	365	ELGYLGSIAICNPVGSSILVRESDLALMTNAGTEIGVASTKFTTQTLTVLLMLVAKLSRLK	424
Db	242	ENGAL-CVGIINTVGTSLRTHKCGVHINAGCEIGVASTKAYTSQIVAMAMMALAIGSDQ	300
Qy	425	GLDASIEHDIVHGLOALPSRIEQMLSQDKRTLEALEDPSDKHALFLSRGDOYPTALEGA	484
Db	301	ISTQARRDSIISGLNNLSSNVSEVLKLDAGMKELASSLIDSESLLVFRGYNVYATALEGA	360
Qy	485	LKLEISYIHAEAAAGELKHGPIALADAMPVTVVAPNNNELLEKLSNIEBVRARGGOL	544
Db	361	LKVKEVALMHSEGMLAGEMKHGPIALVDENLPIIVITRADCFSKQSQSVIQQLLSRGRLL	420
Qy	545	YVFADQ-DAGEV--SSDNMHIENPHVEEVIAPITFVTPLOLLAYHVALIKGTVDVQPRN	601
Db	421	IVMCSRGDASAVCSGGSCRVIYEPVQADCLQPVINIIPLOLLAYHVLTVLRGFDVDQPRN	480
Qy	602	LAKSVTVE 609	
Db	481	LAKSVTTQ 488	
RESULT 15			
US-10-488-015-17			
; Sequence 17, Application US/10488015			
; Publication No. US20060088902A1			
; GENERAL INFORMATION:			
; APPLICANT: STICHTING VOOR DE TECHNISCHE WETENSCHAPPEN			
; APPLICANT: VAN DEN HONDEBL, Cornelis			
; APPLICANT: RAM, Arthur			
; APPLICANT: DAMVELD, Robbert			
; APPLICANT: ARENTSHORST, Mark			
; TITLE OF INVENTION: Methods and Materials for the			
; TITLE OF INVENTION: Identification of Antifungal Substrates in Filamentous Fungi			
; FILE REFERENCE: 13603PCTUS			
; CURRENT APPLICATION NUMBER: US/10/488.015			
; CURRENT FILING DATE: 2004-02-27			
; PRIOR APPLICATION NUMBER: EP 01203423.7			
; PRIOR FILING DATE: 2001-08-28			
; PRIOR APPLICATION NUMBER: PCT/EP02/09639			
; PRIOR FILING DATE: 2002-08-28			
; NUMBER OF SEQ ID NOS: 29			
; SOFTWARE: FastSEQ for Windows Version 4.0			
; SEQ ID NO 17			
; LENGTH: 464			
; TYPE: PRT			
; ORGANISM: Aspergillus niger			
US-10-488-015-17			
Query Match 18.9%; Score 582.5; DB 6; Length 464;			
Best Local Similarity 33.9%; Pred. No. 1.3e-35;			
Matches 156; Conservative 75; Mismatches 152; Indels 77; Gaps 9;			
Qy	96	VHNGIITENHEPLREELKARGYTVSETDTEVIAHLVNWELKQGGTLREAVL--RAIFQL	153
Db	9	VHNGIITNYKELKALKLESFGFETETEDTECIAKLTKLYDQQDPDIDFTVLAKAVVKEL	68
Qy	154	RGAYGTVMDSRHPDTLAARSGSPVLGL-----	183
Db	69	EGAFGLLIKSVHYHVEHVIARAKGSPVLIGVTRSRMKVDFVDVEYSEDGPLPAEQASQNV	128
Qy	184	-----GM-----GENFIASDOLALLPVTRTRFIFLSEG	210
Db	129	AIKKSATGLLAPPDKSLHRSQSRAFLSDDGVPQPAEFLLSSDPSAIVEHTKKVLYLEDD	188
Qy	211	DIABITRRSVNIF----DKTGAEVKQDIESNIQYDAGDKGIYCHYMKQEIYEQPNAIK	265
Db	189	DIAIHGQQLNIHLTKDDGTSNVRAITQTELEQETM--KGNFDHFQKEIFEQESVU	246
Qy	266	NLTGTR--ISHGQVDLSELGFNADLLSKVEHIQILACGTSYNSGMVSRYPFESLAGIPC	323
Db	247	NWEGRLDVANKOVTLLGGLROYI--STTRRCRRRIIFVACGTSYHSCMAVRGVFEELTEIPI	305

```

Qy 324 DVEIASEFRYKSAVRNSLMTLSQGETADTLAGLRLSKELGYLGSLAICNVPGSSLV 383
Db 306 SVELASDFLDQAPVFRDDTCVFSQSGETADSLMALRYCLERGA-TVGIVNVVGSSIS 364
Qy 384 RESDLALMTNAGTEIGVASTKAPTTLTVLLMLVAKLSRLKGLDASIEHDIHVHGLQALPS 443
Db 365 LLTHCGVHINAGPEIGVASTKAYTSQFVAMVPALSSEDRAKQKQKREIMEGLAKVSE 424
Qy 444 RIBQMLSQDKRIEAL-AEDFSDKHHLFLSRGQOYPPIALE 482
Db 425 QFKEILKLNPIQKQCAKFFKNQKSLLLGRCGQOFTALE 464

```

Search completed: June 14, 2006, 15:34:56  
Job time : 22.6172 secs

**THIS PAGE IS BLANK**

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 14, 2006, 15:28:00 ; Search time 57.0938 Seconds  
(without alignments)  
933.660 Million cell updates/sec

Title: US-10-612-779-6  
Perfect score: 3089  
Sequence: 1 MCGIVGAIQRDVAEILLEG.....LIKGTVDQPRNLAKSVTVE 609

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /EMC\_Celerra\_SID33/ptodata/2/iaa/5 COMB.pdp.\*  
2: /EMC\_Celerra\_SID33/ptodata/2/iaa/6 COMB.pdp.\*  
3: /EMC\_Celerra\_SID33/ptodata/2/iaa/7 COMB.pdp.\*  
4: /EMC\_Celerra\_SID33/ptodata/2/iaa/H COMB.pdp.\*  
5: /EMC\_Celerra\_SID33/ptodata/2/iaa/PCTUS COMB.pdp.\*  
6: /EMC\_Celerra\_SID33/ptodata/2/iaa/RE COMB.pdp.\*  
7: /EMC\_Celerra\_SID33/ptodata/2/iaa/backfiles1.pdp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3089	100.0	609	2	US-09-115-475-22
2	3072	99.4	609	2	US-09-115-475-28
3	3072	99.4	609	2	US-09-115-475-31
4	3068	99.3	609	2	US-09-115-475-16
5	3061	99.1	609	2	US-09-115-475-25
6	3053	98.8	609	2	US-09-115-475-19
7	2862	92.7	651	2	US-09-489-039A-10691
8	2498.5	80.9	622	2	US-09-543-681A-7522
9	1983	64.2	616	2	US-09-252-991A-19164
10	1769.5	57.3	613	2	US-09-328-352-5066
11	1491.5	48.3	619	2	US-09-540-236-2182
12	1351.5	43.8	572	2	US-09-302-540-16012
13	1346.5	43.6	592	1	US-08-599-171A-30
14	1346.5	43.6	592	1	US-08-646-590B-30
15	1346.5	43.6	592	2	US-09-069-226-30
16	1346.5	43.6	592	2	US-09-412-184-30
17	1148	37.2	611	2	US-09-438-185A-970
18	1132	36.6	655	2	US-09-107-532A-4425
19	1110	35.9	619	2	US-09-134-001C-4248
20	1106.5	35.8	602	2	US-09-583-110-4085
21	1106.5	35.8	635	2	US-09-107-433-5022
22	1098.5	35.6	682	1	US-08-911-445-3
23	1098.5	35.6	682	2	US-09-182-983-3
24	1098.5	35.6	682	2	US-09-771-838A-3
25	1098.5	35.6	682	2	US-09-949-016-6591
26	1078.5	34.9	680	2	US-09-731-166-6

27	955.5	30.9	615	1	US-08-911-445-2	Sequence 2, Appli
28	955.5	30.9	615	2	US-09-182-983-2	Sequence 2, Appli
29	955.5	30.9	615	2	US-09-771-838A-2	Sequence 2, Appli
30	839.5	27.2	640	2	US-09-134-000C-3871	Sequence 3871, Ap
31	739	23.9	487	2	US-09-248-796A-17250	Sequence 17250, A
32	698	22.6	365	2	US-09-949-016-11214	Sequence 11214, A
33	670	21.7	405	2	US-09-710-279-232	Sequence 232, App
34	605.5	19.6	422	2	US-09-270-767-45187	Sequence 45187, A
35	575	18.6	486	1	US-08-911-445-18	Sequence 18, Appl
36	575	18.6	486	2	US-09-182-983-18	Sequence 18, Appl
37	575	18.6	486	2	US-09-771-838A-18	Sequence 18, Appl
38	574	18.6	425	1	US-08-911-445-1	Sequence 1, Appli
39	574	18.6	425	2	US-09-182-983-1	Sequence 1, Appli
40	574	18.6	425	2	US-09-771-838A-1	Sequence 1, Appli
41	548	17.7	338	2	US-09-710-279-320	Sequence 320, App
42	495	16.0	307	2	US-09-198-452A-1041	Sequence 1041, Ap
43	360.5	11.7	161	2	US-09-198-452A-1040	Sequence 1040, Ap
44	326	10.6	708	2	US-09-252-991A-18074	Sequence 18074, A
45	314	10.2	182	2	US-09-198-452A-1042	Sequence 1042, Ap

## ALIGNMENTS

RESULT 1  
US-09-115-475-22  
; Sequence-22, Application US/09115475  
; Patent No. 6372457  
; GENERAL INFORMATION:  
; APPLICANT: Berry, Alan  
; APPLICANT: Burlingame, Richard P.  
; APPLICANT: Millis, James R.  
; TITLE OF INVENTION: PROCESS AND MATERIALS FOR PRODUCTION OF GLUCOSAMINE  
; FILE REFERENCE: 3161-18-C1  
; CURRENT APPLICATION NUMBER: US/09/115,475  
; EARLIER FILING DATE: 1998-07-15  
; EARLIER APPLICATION NUMBER: PCT/US98/00800  
; EARLIER FILING DATE: 1998-01-14  
; EARLIER APPLICATION NUMBER: 60/035,494  
; EARLIER FILING DATE: 1997-01-14  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 22  
; LENGTH: 609  
; TYPE: PRT  
; ORGANISM: Escherichia coli  
US-09-115-475-22

Query Match	100.0%;	Score 3089;	DB 2;	Length 609;
Best Local Similarity	100.0%;	Pred. No. 3.5e-300;		
Matches 609;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MCGIVGAIQRDVAEILLEGRLRLEVRGYSAGLAVVDTEGHMTRRLRLKQVOMLAQAAE 60		
Db	1	MCGIVGAIQRDVAEILLEGRLRLEVRGYSAGLAVVDTEGHMTRRLRLKQVOMLAQAAE 60		
QY	61	EHPHGGTGIAHTRWATHGEPSEVNAHPVSEHVVVHNGIIEHHPLEELKARGYTFV 120		
Db	61	EHPHGGTGIAHTRWATHGEPSEVNAHPVSEHVVVHNGIIEHHPLEELKARGYTFV 120		
QY	121	SETDTEVIAHLVNWELKQGTLEAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGSPLV 180		
Db	121	SETDTEVIAHLVNWELKQGTLEAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGSPLV 180		
QY	181	IGLGMENFIASQDALLPVTRRFIFLEGDIAEITRRSVNIIDKTAEGVVKRODIESNLQ 240		
Db	181	IGLGMENFIASQDALLPVTRRFIFLEGDIAEITRRSVNIIDKTAEGVVKRODIESNLQ 240		
QY	241	YDAGDKGIYCHYMKQEIYEQPNAIKNTLTGRIISHGVODLSELGNADLLSKVEHIQILA 300		
Db	241	YDAGDKGIYCHYMKQEIYEQPNAIKNTLTGRIISHGVODLSELGNADLLSKVEHIQILA 300		
QY	301	CGTSYNSGMVSRVWFESLAGIPCDVEIASEFRYKSAVRNRNSLMTLSQSGETADTLA 360		

```
Db 301 CGTSYNSGMVSRYPFESLAGIPCDVEIASSEFRYKSAVRNLSMITLSQSGETADTLAGL 360
QY 361 RLSKELGYLGSIAICNVPGSSILVRSDALAMTNAGTEIGVASTKAFITQLTIVLLMLVAKL 420
Db 361 RLSKELGYLGSIAICNVPGSSILVRSDALAMTNAGTEIGVASTKAFITQLTIVLLMLVAKL 420
QY 421 SRLKGLDASIEHDIIVHGLQALPSRIEQMLSQDKRIEALAEFSDKHAFILSRGQOYPIA 480
Db 421 SRLKGLDASIEHDIIVHGLQALPSRIEQMLSQDKRIEALAEFSDKHAFILSRGQOYPIA 480
QY 481 LEGALKLKEISYIHAEAYAAAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEVRAR 540
Db 481 LEGALKLKEISYIHAEAYAAAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEVRAR 540
QY 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPIFYTVPQLLAYHVALIKGTDVQDPR 600
Db 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPIFYTVPQLLAYHVALIKGTDVQDPR 600
QY 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609
```

## RESULT 2

```
US-09-115-475-28
; Sequence 28, Application US/09115475
; Patent No. 6372457
; GENERAL INFORMATION:
; APPLICANT: Berry, Alan
; APPLICANT: Burlingame, Richard P.
; APPLICANT: Millis, James R.
; TITLE OF INVENTION: PROCESS AND MATERIALS FOR PRODUCTION OF GLUCOSAMINE
; FILE REFERENCE: 3161-18-C1
; CURRENT APPLICATION NUMBER: US/09/115,475
; CURRENT FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: PCT/US98/00800
; EARLIER FILING DATE: 1998-01-14
; EARLIER APPLICATION NUMBER: 60/035,494
; EARLIER FILING DATE: 1997-01-14
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-115-475-28
```

```
Query Match 99.4%; Score 3072; DB 2; Length 609;
Best Local Similarity 99.7%; Pred. No. 1.8e-298;
Matches 607; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 MCGIVGATAQRDVAAIILEGLRLRYRGYDSAGLAVVDTGEGHMTLRRLRGKVMQLAAAE 60
Db 1 MCGIVGATAQRDVAAIILEGLRLRYRGYDSAGLAVVDAEGHMTLRRLRGKVMQLAAAE 60
QY 61 EPHLGGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHNEHEPLREELKARGYTFV 120
Db 61 EPHLGGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHNEHEPLREELKARGYTFV 120
QY 121 SETDTEVIAHLNVNELKOGGTLREAVLRAIPOLRGAYGTVMDSRHPDTLLAARSGSPLV 180
Db 121 SETDTEVIAHLNVNELKOGGTLREAVLRAIPOLRGAYGTVMDSRHPDTLLAARSGSPLV 180
QY 181 IGLGMGENFIASDQALLPVTRRRFIFLEEGDIAETTRRSVNI FDKTGAEVKQDIESNIQ 240
Db 181 IGLGMGENFIASDQALLPVTRRRFIFLEEGDIAETTRRSVNI FDKTGAEVKQDIESNIQ 240
QY 241 YDAGDKGIYCHYMOKEIEYEQPNAIKNTLTGRI SHGQVDLSSELGPNADELLSKVEHIQILA 300
Db 241 YDAGDKGIYCHYMOKEIEYEQPNAIKNTLTGRI SHGQVDLSSELGPNADELLSKVEHIQILA 300
QY 301 CGTSYNSGMVSRYPFESLAGIPCDVEIASSEFRYKSAVRNLSMITLSQSGETADTLAGL 360
```

```
Db 301 CGTSYNSGMVSRYPFESLAGIPCDVEIASSEFRYKSAVRNLSMITLSQSGETADTLAGL 360
QY 361 RLSKELGYLGSIAICNVPGSSILVRSDALAMTNAGTEIGVASTKAFITQLTIVLLMLVAKL 420
Db 361 RLSKELGYLGSIAICNVPGSSILVRSDALAMTNAGTEIGVASTKAFITQLTIVLLMLVAKL 420
QY 421 SRLKGLDASIEHDIIVHGLQALPSRIEQMLSQDKRIEALAEFSDKHAFILSRGQOYPIA 480
Db 421 SRLKGLDASIEHDIIVHGLQALPSRIEQMLSQDKRIEALAEFSDKHAFILSRGQOYPIA 480
QY 481 LEGALKLKEISYIHAEAYAAAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEVRAR 540
Db 481 LEGALKLKEISYIHAEAYAAAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEVRAR 540
QY 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPIFYTVPQLLAYHVALIKGTDVQDPR 600
Db 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPIFYTVPQLLAYHVALIKGTDVQDPR 600
QY 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609
```

## RESULT 3

```
US-09-115-475-31
; Sequence 31, Application US/09115475
; Patent No. 6372457
; GENERAL INFORMATION:
; APPLICANT: Berry, Alan
; APPLICANT: Burlingame, Richard P.
; APPLICANT: Millis, James R.
; TITLE OF INVENTION: PROCESS AND MATERIALS FOR PRODUCTION OF GLUCOSAMINE
; FILE REFERENCE: 3161-18-C1
; CURRENT APPLICATION NUMBER: US/09/115,475
; CURRENT FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: PCT/US98/00800
; EARLIER FILING DATE: 1998-01-14
; EARLIER APPLICATION NUMBER: 60/035,494
; EARLIER FILING DATE: 1997-01-14
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-115-475-31
```

```
Query Match 99.4%; Score 3072; DB 2; Length 609;
Best Local Similarity 99.7%; Pred. No. 1.8e-298;
Matches 607; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 MCGIVGATAQRDVAAIILEGLRLRYRGYDSAGLAVVDTGEGHMTLRRLRGKVMQLAAAE 60
Db 1 MCGIVGATAQRDVAAIILEGLRLRYRGYDSAGLAVVDAEGHMTLRRLRGKVMQLAAAE 60
QY 61 EPHLGGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHNEHEPLREELKARGYTFV 120
Db 61 EPHLGGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHNEHEPLREELKARGYTFV 120
QY 121 SETDTEVIAHLNVNELKOGGTLREAVLRAIPOLRGAYGTVMDSRHPDTLLAARSGSPLV 180
Db 121 SETDTEVIAHLNVNELKOGGTLREAVLRAIPOLRGAYGTVMDSRHPDTLLAARSGSPLV 180
QY 181 IGLGMGENFIASDQALLPVTRRRFIFLEEGDIAETTRRSVNI FDKTGAEVKQDIESNIQ 240
Db 181 IGLGMGENFIASDQALLPVTRRRFIFLEEGDIAETTRRSVNI FDKTGAEVKQDIESNIQ 240
QY 241 YDAGDKGIYCHYMOKEIEYEQPNAIKNTLTGRI SHGQVDLSSELGPNADELLSKVEHIQILA 300
Db 241 YDAGDKGIYCHYMOKEIEYEQPNAIKNTLTGRI SHGQVDLSSELGPNADELLSKVEHIQILA 300
QY 301 CGTSYNSGMVSRYPFESLAGIPCDVEIASSEFRYKSAVRNLSMITLSQSGETADTLAGL 360
```



Db 301 CGTSYNGMVSRYWFSFSLAGIPCDVEIASSEFRYKSAVRNSLMTLSQSGETADTLA 360  
QY 361 RLSKELGYLGSLAICNVPSSLVRESDLALMTNAGTEIGVASTKAFTTQTLTVLLMLVAKL 420  
Db 361 RLSKELGYLGSLAICNVPSSLVRESDLALMTNAGTEIGVASTKAFTTQTLTVLLMLVAKL 420  
QY 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEEDFSDKHAFSLRGDOYPIA 480  
Db 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEEDFSDKHAFSLRGDOYPIA 480  
QY 481 LEGALKLKEISYTHAEAYAGELKHGFLALIDAMPVIVVAPNNLEKLSNIEEVRAR 540  
Db 481 LEGALKLKEISYTHAEAYAGELKHGFLALIDAMPVIVVAPNNLEKLSNIEEVRAR 540  
QY 541 GGQLYVFADQDAGFVSSDNHIIEMPHVEEVIAPFYTVPLQLLAYHVALIKGTDVQPR 600  
Db 541 GGQLYVFADQDAGFVSSDNHIIEMPHVEEVIAPFYTVPLQLLAYHVALIKGTDVQPR 600  
QY 601 NLAKESTVE 609  
Db 601 NLAKESTVE 609

## RESULT 4

US-09-115-475-16  
; Sequence 16, Application US/09115475  
; Patent No. 6372457  
; GENERAL INFORMATION:  
; APPLICANT: Berry, Alan  
; APPLICANT: Burlingame, Richard P.  
; APPLICANT: Millis, James R.  
; TITLE OF INVENTION: PROCESS AND MATERIALS FOR PRODUCTION OF GLUCOSAMINE  
; FILE REFERENCE: 3161-18-C1  
; CURRENT APPLICATION NUMBER: US/09/115,475  
; CURRENT FILING DATE: 1998-07-15  
; EARLIER APPLICATION NUMBER: PCT/US98/00800  
; EARLIER FILING DATE: 1998-01-14  
; EARLIER APPLICATION NUMBER: 60/035,494  
; EARLIER FILING DATE: 1997-01-14  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 16  
; LENGTH: 609  
; TYPE: PRT  
; ORGANISM: Escherichia coli  
US-09-115-475-16

Query Match 99.3%; Score 3068; DB 2; Length 609;  
Best Local Similarity 99.5%; Pred. No. 4.4e-298;  
Matches 606; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 MCGIVGAIQAQDVAEILLLEGLRLLEYRGYDSAGLAVVDTEGHMTRRLRLGKVQMLQAQAE 60  
Db 1 MCGIVGAIQAQDVAEILLLEGLRLLEYRGYDSAGLAVVDTEGHMTRRLRLGKVQMLQAQAE 60  
QY 61 EPHLHGCTGTAHTRWATHGEPSEVNAHPHVSHEIIVVHNGIIEHHEPLREELKARGYTFV 120  
Db 61 EPHLHGCTGTAHTRWATHGEPSEVNAHPHVSHEIIVVHNGIIEHHEPLREELKARGYTFV 120  
QY 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGSPV 180  
Db 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGSPV 180  
QY 181 IGLGMENFIASDQALLPVTRRFIFLEEGDIAEITRRSVNIIPDKTGAEVKRODIESNLQ 240  
Db 181 IGLGMENFIASDQALLPVTRRFIFLEEGDIAEITRRSVNIIPDKTGAEVKRODIESNLQ 240  
QY 241 YDAGDKGIYCHYMQKEIYEOPNAIKNTLTGRISHGQVDLSLGNADLLEKLSKVEHIQILA 300  
Db 241 YDAGDKGIYCHYMQKEIYEOPNAIKNTLTGRISHGQVDLSLGNADLLEKLSKVEHIQILA 300  
QY 301 CGTSYNGMVSRYWFSFSLAGIPCDVEIASSEFRYKSAVRNSLMTLSQSGETADTLA 360

Db 301 CGTSYNGMVSRYWFSFSLAGIPCDVEIASSEFRYKSAVRNSLMTLSQSGETADTLA 360  
QY 361 RLSKELGYLGSLAICNVPSSLVRESDLALMTNAGTEIGVASTKAFTTQTLTVLLMLVAKL 420  
Db 361 RLSKELGYLGSLAICNVPSSLVRESDLALMTNAGTEIGVASTKAFTTQTLTVLLMLVAKL 420  
QY 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEEDFSDKHAFSLRGDOYPIA 480  
Db 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEEDFSDKHAFSLRGDOYPIA 480  
QY 481 LEGALKLKEISYTHAEAYAGELKHGFLALIDAMPVIVVAPNNLEKLSNIEEVRAR 540  
Db 481 LEGALKLKEISYTHAEAYAGELKHGFLALIDAMPVIVVAPNNLEKLSNIEEVRAR 540  
QY 541 GGQLYVFADQDAGFVSSDNHIIEMPHVEEVIAPFYTVPLQLLAYHVALIKGTDVQPR 600  
Db 541 GGQLYVFADQDAGFVSSDNHIIEMPHVEEVIAPFYTVPLQLLAYHVALIKGTDVQPR 600  
QY 601 NLAKESTVE 609  
Db 601 NLAKESTVE 609

## RESULT 5

US-09-115-475-25  
; Sequence 25, Application US/09115475  
; Patent No. 6372457  
; GENERAL INFORMATION:  
; APPLICANT: Berry, Alan  
; APPLICANT: Burlingame, Richard P.  
; APPLICANT: Millis, James R.  
; TITLE OF INVENTION: PROCESS AND MATERIALS FOR PRODUCTION OF GLUCOSAMINE  
; FILE REFERENCE: 3161-18-C1  
; CURRENT APPLICATION NUMBER: US/09/115,475  
; CURRENT FILING DATE: 1998-07-15  
; EARLIER APPLICATION NUMBER: PCT/US98/00800  
; EARLIER FILING DATE: 1998-01-14  
; EARLIER APPLICATION NUMBER: 60/035,494  
; EARLIER FILING DATE: 1997-01-14  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 25  
; LENGTH: 609  
; TYPE: PRT  
; ORGANISM: Escherichia coli  
US-09-115-475-25

Query Match 99.1%; Score 3061; DB 2; Length 609;  
Best Local Similarity 99.3%; Pred. No. 2.2e-297;  
Matches 605; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 1 MCGIVGAIQAQDVAEILLLEGLRLLEYRGYDSAGLAVVDTEGHMTRRLRLGKVQMLQAQAE 60  
Db 1 MCGIVGAIQAQDVAEILLLEGLRLLEYRGYDSAGLAVVDTEGHMTRRLRLGKVQMLQAQAE 60  
QY 61 EPHLHGCTGTAHTRWATHGEPSEVNAHPHVSHEIIVVHNGIIEHHEPLREELKARGYTFV 120  
Db 61 EPHLHGCTGTAHTRWATHGEPSEVNAHPHVSHEIIVVHNGIIEHHEPLREELKARGYTFV 120  
QY 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGSPV 180  
Db 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGSPV 180  
QY 181 IGLGMENFIASDQALLPVTRRFIFLEEGDIAEITRRSVNIIPDKTGAEVKRODIESNLQ 240  
Db 181 IGLGMENFIASDQALLPVTRRFIFLEEGDIAEITRRSVNIIPDKTGAEVKRODIESNLQ 240  
QY 241 YDAGDKGIYCHYMQKEIYEOPNAIKNTLTGRISHGQVDLSLGNADLLEKLSKVEHIQILA 300  
Db 241 YDAGDKGIYCHYMQKEIYEOPNAIKNTLTGRISHGQVDLSLGNADLLEKLSKVEHIQILA 300  
QY 301 CGTSYNGMVSRYWFSFSLAGIPCDVEIASSEFRYKSAVRNSLMTLSQSGETADTLA 360

```
Db 301 CGTSYNSGMVSRWFESLAGIPCDVEIASFEFRYKSAVRNLSMTLTSQSGETADTLAGL 360
Qy 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKATFTTQLTVLLMLVAKL 420
Db 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKATFTTQLTVLLMLVAKL 420
Qy 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSQDKRIEALAEAFSDKHAFLSRGDQYPTA 480
Db 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSQDKRIEALAEAFSDKHAFLSRGDQYPTA 480
Qy 481 LEGALKLKEISYIHAEYAAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEEVRAR 540
Db 481 LEGALKLKEISYIHAEYAAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEEVRAR 540
Qy 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPIFYTVPQLLAYHVALIKGTVDQDPR 600
Db 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPIFYTVPQLLAYHVALIKGTVDQDPR 600
Qy 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609

RESULT 6
US-09-115-475-19
; Sequence 19, Application US/09115475
; Patent No. 6372457
; GENERAL INFORMATION:
; APPLICANT: Berry, Alan
; APPLICANT: Burlingame, Richard P.
; APPLICANT: Mills, James R.
; TITLE OF INVENTION: PROCESS AND MATERIALS FOR PRODUCTION OF GLUCOSAMINE
; FILE REFERENCE: 3161-18-C1
; CURRENT APPLICATION NUMBER: US/09/115,475
; CURRENT FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: PCT/US98/00800
; EARLIER FILING DATE: 1998-01-14
; EARLIER APPLICATION NUMBER: 60/035,494
; EARLIER FILING DATE: 1997-01-14
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-115-475-19

Query Match 98.8%; Score 3053; DB 2; Length 609;
Best Local Similarity 99.0%; Pred. No. 1.4e-296;
Matches 603; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
Qy 1 MCGIVGAIAQRDVAEILLEGRLRLEYRGYDSAGLAVVDTEGHMTLRRLRGKQVQLAAQAE 60
Db 1 MCGTVGAIAQRDVAEILLEGRLRLEYRGYDSAGLAVVDAEGHMTLRRLRGKQVQLAAQAE 60
Qy 61 EPHLGGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHNEHPELREELKARGYTFV 120
Db 61 EPHLGGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHNEHPELREELKARGYTFV 120
Qy 121 SETDTEVIAHLVNNELKOGGTLREAVLRAIPQLRGAYGTVIMDSRHPDTLLAARSGSPV 180
Db 121 SETDTEVIAHLVNNELKOGGTLREAVLRAIPQLRGAYGTVIMDSRHPDTLLAARSGSPV 180
Qy 181 IGLGNGENFIASDQALLPVTRRFIFLEEGDIAETTRRSVNI FDKTGAEVKQDIESNLQ 240
Db 181 IGLGNGENFIASDQALLPVTRRFIFLEEGDIAETTRRSVNI FDKTGAEVKQDIESNLQ 240
Qy 241 YDAGDKGIYCHYMQKEIYEQPNAINKNTLTGRISHGQVDLSSELGNADLLSKVEHIQILA 300
Db 241 YDAGDKGIYCHYMQKEIYEQPNAINKNTLTGRISHGQVDLSSELGNADLLSKVEHIQILA 300
Qy 301 CGTSYNSGMVSRWFESLAGIPCDVEIASFEFRYKSAVRNLSMTLTSQSGETADTLAGL 360
Db 301 CGTSYNSGMVSRWFESLAGIPCDVEIASFEFRYKSAVRNLSMTLTSQSGETADTLAGL 360
```

```
Db 301 CGTSYNSGMVSRWFESLAGIPCDVEIASFEFRYKSAVRNLSMTLTSQSGETADTLAGL 360
Qy 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKATFTTQLTVLLMLVAKL 420
Db 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKATFTTQLTVLLMLVAKL 420
Qy 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSQDKRIEALAEAFSDKHAFLSRGDQYPTA 480
Db 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSQDKRIEALAEAFSDKHAFLSRGDQYPTA 480
Qy 481 LEGALKLKEISYIHAEYAAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEEVRAR 540
Db 481 LEGALKLKEISYIHAEYAAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEEVRAR 540
Qy 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPIFYTVPQLLAYHVALIKGTVDQDPR 600
Db 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPIFYTVPQLLAYHVALIKGTVDQDPR 600
Qy 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609

RESULT 7
US-09-489-039A-10691
; Sequence 10691, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10691
; LENGTH: 651
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10691

Query Match 92.7%; Score 2862; DB 2; Length 651;
Best Local Similarity 91.8%; Pred. No. 2.2e-277;
Matches 559; Conservative 29; Mismatches 21; Indels 0; Gaps 0;
```

```
Qy 1 MCGIVGAIAQRDVAEILLEGRLRLEYRGYDSAGLAVVDTEGHMTLRRLRGKQVQLAAQAE 60
Db 43 MCGIVGAIAQRDVAEILLEGRLRLEYRGYDSAGLAVVDSEGHMTVRRLRGKQVQLAAQAE 102
Qy 61 EPHLGGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHNEHPELREELKARGYTFV 120
Db 103 EPHLGGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHNEHPELREELKARGYTFV 162
Qy 121 SETDTEVIAHLVNNELKOGGTLREAVLRAIPQLRGAYGTVIMDSRHPDTLLAARSGSPV 180
Db 163 TETDTEVIAHLVHVELEQGGTLREAVLRAIPQLRGAYGTVIMDTRDPGTLAARSGSPV 222
Qy 181 IGLGNGENFIASDQALLPVTRRFIFLEEGDIAETTRRSVNI FDKTGAEVKQDIESNLQ 240
Db 223 IGLGNGENFIASDQALLPVTRRFIFLEEGDIAETTRRSVNI FDKSAQVKKQEIENSLQ 282
Qy 241 YDAGDKGIYCHYMQKEIYEQPNAINKNTLTGRISHGQVDLSSELGNADLLSKVEHIQILA 300
Db 283 YDAGDKGIYCHYMQKEIYEQPNAINKNTLTGRISHGQVDLSSELGNADLLSKVEHIQILA 342
Qy 301 CGTSYNSGMVSRWFESLAGIPCDVEIASFEFRYKSAVRNLSMTLTSQSGETADTLAGL 360
Db 343 CGTSYNSGMVSRWFESLAGIPCDVEIASFEFRYKSAVRNLSMTLTSQSGETADTLAGL 402
Qy 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKATFTTQLTVLLMLVAKL 420
```

Db 403 RLKELGSLALCNVPGSSLVRESLALMTKAGTEIGVASTKFTTQLTVLLMLVAKL 462  
Qy 421 SRLKGLDASIEDHIVHGLQALPSRIEQLMSQDKRIEALAEFSDKKHALFLSRGQDYPIA 480  
Db 463 ARLKGDASIEDHIVHGLQALPNRIEQLMSQDKRIEQLAEFSDKKHALFLSRGQDYPIA 522  
Qy 481 LEGALKLKEISYTHAEYAAAGELKHGPLALIDAMPVIVVAPNNELLEKLKSNIEEVRAR 540  
Db 523 MEGALKLKEISYTHAEYAAAGELKHGPLALIDAMPVIVVAPNNELLEKLKSNIEEVRAR 582  
Qy 541 GGOLYVPADQDAGFVSDNNHIIEMPHVEEVIAPFTVPLQLLAYHVALIKGTDVDPQR 600  
Db 583 GGELYVPADGEAGFNGSDNNHIIEMPHVEEVIAPFTVPLQLLAYHVALIKGTDVDPQR 642  
Qy 601 NLAKSVTVE 609  
Db 643 NLAKSVTVE 651

RESULT 8  
US-09-543-681A-7522  
; Sequence 7522, Application US/09543681A  
; Patent No. 6605709  
; GENERAL INFORMATION:  
; APPLICANT: GARY BRETON  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
; FILE REFERENCE: 2709.1002-001  
; CURRENT APPLICATION NUMBER: US/09/543.681A  
; PRIOR FILING DATE: 2000-04-05  
; PRIOR FILING DATE: 1999-04-09  
; NUMBER OF SEQ ID NOS: 8344  
; SEQ ID NO 7522  
; LENGTH: 622  
; TYPE: PRT  
; ORGANISM: Proteus mirabilis  
US-09-543-681A-7522

Query Match 80.9%; Score 2498.5; DB 2; Length 622;  
Best Local Similarity 80.8%; Pred. No. 5.6e-241;  
Matches 492; Conservative 46; Mismatches 70; Indels 1; Gaps 1;

Qy 1 MCGIVGAIAORDVAEILLEGRLRLEYRGYDSAGLAVVDTEGHMTRRLRLRGKVQMLAAAE 60  
Db 15 MCGIVGAIAORDVAEILLEGRLRLEYRGYDSAGLAVVDNDCHMTRRLRGKVQMLAAEAE 74  
Qy 61 EPHLGGTGTAHTRWATHGPPSEVNAHPVSEHIVVHNGIIEHNEHPLREELKARGYTFV 120  
Db 75 KTVQVIGTGTIAHTRWATHGPPCQDNAPHVSGTIAVVHNGIIEHNEHPLREELKARGYTFV 134  
Qy 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGPLV 180  
Db 135 SQDTEVIAHLVNWELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGPLV 194  
Qy 181 IGLMGENFTASDQALALLPVTRRPFIEEGDIAETRISVNIIFDKTGAEVKRODIESNLQ 240  
Db 195 VGLGVGENFLASDQALALLPVTRRPFIEEGDIAETRISVNIIFDKTGAEVKRODIESNLQ 254  
Qy 241 YDAGDKGIYCHYMQKEIYEQPNAIKNTLTGRI SHGQVDLSLSELGNADLLSKVEHIQILA 300  
Db 255 YDAGDKGIYCHYMQKEIYEQPNAIKNTLTGRI SHGQVDLSLSELGNADLLSKVEHIQILA 314  
Qy 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASEFPRYKSAVRNNSLMTLSQSGETADTLAAGL 360  
Db 315 CGTSYNSGMVSRVWFESLAGIPCDVEIASEFPRYKSAVRNNSLMTLSQSGETADTLAAGL 374  
Qy 361 RLKELGSLALCNVPGSSLVRESLALMTNAGTEIGVASTKFTTQLTVLLMLVAKL 420  
Db 375 RLKELGSLALCNVAGSSLVRESFVLMTKAGBIGVASTKFTTQLAVLLMLVAYM 434  
Qy 421 SRLKGLDASIEDHIVHGLQALPSRIEQLMSQDKRIEALAEFSDKKHALFLSRGQDYPIA 480

Db 435 GRKGVEA-LEQVAMALHALPSRIESMLSKDVIEALAEFSEKSHALFLRGDQYPIA 493  
Qy 481 LEGALKLKEISYTHAEYAAAGELKHGPLALIDAMPVIVVAPNNELLEKLKSNIEEVRAR 540  
Db 494 VEGALKLKEISYTHAEYAAAGELKHGPLALIDAMPVIVVAPNNELLEKLKSNIEEVRAR 553  
Qy 541 GGOLYVPADQDAGFVSDNNHIIEMPHVEEVIAPFTVPLQLLAYHVALIKGTDVDPQR 600  
Db 554 GGELYVPADQDAGFESETWKLISLPHVEELIAPFTVPLQLLSYHVALIKGTDVDPQR 613  
Qy 601 NLAKSVTVE 609  
Db 614 NLAKSVTVE 622

RESULT 9  
US-09-252-991A-19164  
; Sequence 19164, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252.991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 19164  
; LENGTH: 616  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-19164

Query Match 64.2%; Score 1983; DB 2; Length 616;  
Best Local Similarity 62.4%; Pred. No. 2.6e-189;  
Matches 381; Conservative 101; Mismatches 127; Indels 2; Gaps 2;

Qy 1 MCGIVGAIAORDVAEILLEGRLRLEYRGYDSAGLAVVDTEGHMTRRLRLRGKVQMLAAAE 60  
Db 6 MCGIVGAIAERNTPILIEGLKRLLEYRGYDSAGVAVFDNEGRLQRCRVKVASLEGLA 65  
Qy 61 EPHLGGTGTAHTRWATHGPPSEVNAHPV-SEHIVVHNGIIEHNEHPLREELKARGYTF 119  
Db 66 GTELLGRLGIAHTRWATHGAPTEGNAHPFSSDELAHVHNGIIEHNEHPLREELKGLGYF 125  
Qy 120 VSETDTEVIAHLVNWELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGPL 179  
Db 126 TSQDTEVIVHLLHKLQSIGDITLAKDAVKELHGLAYGLAVISAAQPDRIVAARSGPL 185  
Qy 180 VIGLGVGENFIASDQALALLPVTRRPFIEEGDIAETRISVNIIFDKTGAEVKRODIESNL 239  
Db 186 VIGLGVGENFIASDQALALLPVTRRPFIEEGDIAETRISVNIIFDKTGAEVKRODIESNL 245  
Qy 240 QYDAGDKGIYCHYMQKEIYEQPNAIKNTLTGRI SHGQVDLSLSELGNADLLSKVEHIQIL 299  
Db 246 GAERADKGEYRHFMKLEIHEQPSVVQRTLEGRLQGVVMEVSTGPOAELFAKVRNVQIV 305  
Qy 300 ACCTSNGMVSRVWFESLAGIPCDVEIASEFPRYKSAVRNNSLMTLSQSGETADTLAAGL 359  
Db 306 ACCTSNGMVSRVWFESLAGIPCDVEIASEFPRYKSAVRNNSLMTLSQSGETADTLAAGL 365  
Qy 360 LRLSKELGSLALCNVPGSSLVRESLALMTNAGTEIGVASTKFTTQLTVLLMLVAK 419  
Db 366 LRNAELGFLUSSVAICNVATSSLVRESLALMTNAGTEIGVASTKFTTQLTVLLMLVAK 425  
Qy 420 LSRLL-KGLDASIEDHIVHGLQALPSRIEQLMSQDKRIEALAEFSDKKHALFLSRGQY 478  
Db 426 IGQVQKELADGVEAELVDELRLRLTRLEALAMNRITVEKVSSELFAEKHHTFLGRGAQFP 485





```

Db 405 ALSVR-----ESERENLRLLKVPVSLVQTLTAAEEVKVAEKYMKKQNMVLYLGRY 457
QY 475 DOYPALGALKLKEISYIHAAYAGELKHGFLALIDADMPVIVVAPNNLELKLKSN1 534
Db 458 LNPYALGALKLKEISYIHAEGYPAGEMKHGPIALIDENMPVVVIAPKDRVYEKILSNV 517
QY 535 BEVRARGQLVYFADQDAGFVSSDNMHIEMPHVEEVIAPFYTPVLPOLLAYHVALIKGT 594
Db 518 BEVLARKGRVISVGFKGDDETLKSKSESVMIEPKAEEPTTPELTVIPLQLFAYFIASKLGL 577
QY 595 DVDQPRNLAKSVTVE 609
Db 578 DVDQPRNLAKTVTVE 592

RESULT 14
US-08-646-590B-30
; Sequence 30, Application US/08646590B
; Patent No. 5962283
; GENERAL INFORMATION:
; APPLICANT: Warren, Patrick V.
; APPLICANT: Swanson, Ronald V.
; TITLE OF INVENTION: TRANSAMINASES AND AMINOTRANSFERASES
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: Fast-SEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,590B
; FILING DATE: 08-May-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/599,171
; FILING DATE: 09-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/01094
; FILING DATE: 21-January-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Ph.D., Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 09010/017001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 592 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-646-590B-30

Query Match 43.6%; Score 1346.5; DB 1; Length 592;
Best Local Similarity 46.3%; Pred. No. 1.5e-125;
Matches 285; Conservative 112; Mismatches 189; Indels 29; Gaps 9;

QY 1 MCGIVGATAQRDVA-EILLEGRLRLRYEGYDSAGLAVVDTEGHMTRRLRLGKVMQAQA 59
Db 1 MCGIVGVYG-RDLALPVLGALERLEYRGYDSAGVALIE-DGKLIVBKKGKIRELVKAL 58
QY 60 BEHPLHGHTGIAHTRWATHGEBSEVNAHPHVE--HIVVHNGIENHEPLREELKARGY 117

```

```

Db 59 WGYKAKTGTHRWATHGRTDENAHPHDEKGEFAVVRHNGIENLELEKEELKKEGV 118
QY 118 TFSVETDTEVIAHLVNWELKQGGTLREAVLRAIPQIRGAYGTIVIMDSRHPDTLAARSGS 177
Db 119 KFRSETDTEVIAHLIAKNYR--GDLEAVLTKVKLKGAFAPAVITVHEPNRLIGVKQGS 176
QY 178 PLVITGLMGNPIASDQJALLPVTRRFIFLEGDIAETIRRSVNIFFDKTGABVKKQDIES 237
Db 177 PLIVLGBGENFLASDIPAILPYTKIIVLDDGETADLTPTDVTNIYNFEGEPVSEVMIT 236
QY 238 NLOVDAGDKGIYCHVMOKEIYEQPNAIKNTLTGRTISHGQVDLSSELGPNADEL--LSKVE 294
Db 237 PWDLVSAEKGKGFHMLKEIYEQPKAINDTLKGFLS-----TEDAIPFKLKDPR 285
QY 295 HIQILACGTSYNSGMVSRVWFESLAGIPCDVEIASFEFRRKSAVRNRNLSMITLSQGETA 354
Db 286 RVLIACGTSYHAGFVGKWIETRFAGVPTVEIYASEFRYADVPSDKDIVIGISQGETA 345
QY 355 DTLAGRLSKELGYLGSIAICNVPGSSLVRSDDLALMTNAGTEIGVASTKATTTQLTULL 414
Db 346 DTKFALQSAKEKGF-TVGLVNVVGSADRESDFSLHTHAGEIIGVAATKTTTAQFTALY 404
QY 415 MLVAKLSRLKGLDASIEHDIYHGLQALPSRIEQMLSQDKRIEALAEFSDKHHLFLSRG 474
Db 405 ALSVR-----ESEERENLRLLKVPVSLVQTLTAAEEVKVAEKYMKKQNMVLYLGRY 457
QY 475 DOYPALGALKLKEISYIHAAYAGELKHGFLALIDADMPVIVVAPNNLELKLKSN1 534
Db 458 LNPYALGALKLKEISYIHAEGYPAGEMKHGPIALIDENMPVVVIAPKDRVYEKILSNV 517
QY 535 BEVRARGQLVYFADQDAGFVSSDNMHIEMPHVEEVIAPFYTPVLPOLLAYHVALIKGT 594
Db 518 BEVLARKGRVISVGFKGDDETLKSKSESVMIEPKAEEPTTPELTVIPLQLFAYFIASKLGL 577
QY 595 DVDQPRNLAKSVTVE 609
Db 578 DVDQPRNLAKTVTVE 592

RESULT 15
US-09-069-226-30
; Sequence 30, Application US/09069226
; Patent No. 6013509
; GENERAL INFORMATION:
; APPLICANT: WARREN, Patrick V.
; TITLE OF INVENTION: TRANSAMINASES AND AMINOTRANSFERASES
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA: US/09/069,226
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/599,171
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: HERRON, CHARLES J.
; REGISTRATION NUMBER: 28,019
; REFERENCE/DOCKET NUMBER: 331400-38
; TELECOMMUNICATION INFORMATION:

```

```
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 592 AMINO ACIDS
;   TYPE: AMINO ACID
;   TOPOLOGY: LINEAR
;   MOLECULE TYPE: PROTEIN
;   US-09-069-226-30

Query Match      43.6%; Score 1346.5; DB 2; Length 592;
Best Local Similarity 46.3%; Pred. No. 1.5e-125;
Matches 285; Conservative 112; Mismatches 189; Indels 29; Gaps 9;

QY      1 MCGIQAIAQRDVA-EILLEGRLRYRGYDSAGLAVVDTEGHMTRLRLRGKVMQLAQAA 59
Db      1 MCGIVGVG-RDLALPIVLGALERLEYRGYDSAGVALIE-DGKLIIVEKKGKIRELVKAL 58

QY      60 EEHPLHGGTGIATRWATHGEPSEVNAHPHVS--HIVVVHNGIIEHNEPLREELKARGY 117
Db      59 WGDYKAKTGICHTRWATHGKPTDENAHPTDEKGEPAVVHNGIIEHNEPLREELKKEGV 118

QY      118 TFVSETDTEVIAHLVNMWELKOGTLEAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGS 177
Db      119 KFRSETDTEVIAHLIAKNR--GDLLEAVLTKVKLKGAFAFAVITVHEPNRLIGVKQGS 176

QY      178 PLVIGLGMENFIASDQALLPVTTRRFIFLEEGDIAEITRRSVNIFDKTGAEVKQRDIES 237
Db      177 PLIVGLGEGENFLASDIPAILPYTKIIVLDGGEIADLTPTDVTNIVNFEGEPSKEVMIT 236

QY      238 NLQYDAGDKGIYCHYMOKEIYEQPNAIKNTLTGRISHGQVDLSLGNADL---LSKVE 294
Db      237 PWDLVSAEKGFGFHMKEIYEQPKAINDTLKGFLS-----TEDAIFPKLKDFR 285

QY      295 HIQILACGTSYNSGWSRYWFESLAGIPCDVEIASEFRYKRSVRRNSLMTLSQGETA 354
Db      286 RVLIACGTSYHAGFVGKYWIERFAGVPTVIVASEPRYADVPVSDKDIVIGISQGETA 345

QY      355 DTLAGRLSKELGYLSLAICNVPSSLVRESDLALMTNAGTEIGVASTKAFITOLTVLL 414
Db      346 DTKFALQSAKEGAF-TVGLVNVVGSADRESDFSLHTHAGPEIGVAATKTFTAQFTALY 404

QY      415 MLVAKLSRLKGLDASIEHDIHGLQALPSRIQWLSQDKRIEALAEDEFDKHHALFLSRG 474
Db      405 ALSVR-----ESERENILIRLEKVPISLVEQTLNTAEVEKVAEKYMKKKNNWLYLGRY 457

QY      475 DQYPIALEGALKLKEISYTHAEAYAGELKHGPIALIDADMPVIVVAPNNELLEKLKSN 534
Db      458 LNYPIALEGALKLKEISYTHAEGYPAGEMKHGPIALIDENMPVVVIAPKDRVVEKILSNV 517

QY      535 EEVRARGGQLYVFADQDAGVFSSDNMHIIEMPHVEEVIAPIFTVPLQLLAYHVALIKGT 594
Db      518 EEVLARKGRVISVFGKDETLKSKSESVMBIPKAEETITPFLTVPILQLFAYFIASKLGL 577

QY      595 DVDQPRNLAKSVTVE 609
Db      578 DVDQPRNLAKTIVE 592
```

Search completed: June 14, 2006, 15:29:54  
Job time : 59.0938 secs

**THIS PAGE IS BLANK**



GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 14, 2006, 15:15:09 ; Search time 305.293 Seconds  
(without alignments)  
1845.226 Million cell updates/sec

Title: US-10-612-779-6  
Perfect score: 3089  
Sequence: 1 MCGIVGAIQRDAEILLEG.....LIKGTVDQPRNLAKSVTVE 609

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 7.2.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result	No.	Score	Query Match	Length	DB	ID	Description
1	3068	99.3	609	2	Q2M847	ECOLI	Q2M847 escherichia
2	3063	99.2	608	1	GLMS	ECOLI	P17169 e glucosami
3	3063	99.2	608	1	Q329R8	SHIDS	Q329R8 shigella dy
4	3061	99.1	609	2	Q31UM9	SHIDS	Q31UM9 shigella bo
5	3059	99.0	608	1	GLMS	ECOS7	Q8xeg2 e glucosami
6	3059	99.0	608	1	GLMS	SHIFL	Q83iy4 s glucosami
7	3055	98.9	609	2	Q3YVN3	SHIDS	Q3YVN3 shigella so
8	3049	98.7	608	1	GLMS	ECOL6	Q8ftb4 e glucosami
9	3040	98.4	608	1	GLMS	SALPA	Q5pkv9 s glucosami
10	3040	98.4	609	2	Q57HY2	SALCH	Q57HY2 salmonella
11	3037	98.3	608	1	GLMS	SALTY	Q8z2q2 s glucosami
12	3030	98.1	608	1	GLMS	SALTI	Q8z2q2 s glucosami
13	2667	86.4	609	1	GLMS	ERWCT	Q8cyj9 e glucosami
14	2663	86.2	608	1	GLMS	YERPE	Q8z9s8 y glucosami
15	2663	86.2	608	1	GLMS	YERPS	Q8z9s8 y glucosami
16	2553	82.6	608	1	GLMS	PHOLL	Q7na97 p glucosami
17	2532	82.0	610	2	Q2N083	SODGL	Q2Nq83 sodalis glo
18	2260	73.2	610	2	Q4QNB6	HAET8	Q4qnb6 haemophilus
19	2251	72.9	609	1	GLMS	HA8IN	P44708 h glucosami
20	2245	72.7	609	1	GLMS	VIBCH	Q9kun8 v glucosami
21	2243	72.6	609	1	GLMS	VIBF1	Q8e279 v glucosami
22	2236	72.4	609	1	GLMS	PHOPR	Q61lh3 p glucosami
23	2214	71.7	609	1	GLMS	VIBPA	Q87sr3 v glucosami
24	2210	71.6	609	1	GLMS	VIBVU	Q8def3 v glucosami
25	2208	71.5	609	1	GLMS	VIBVY	Q7mp62 v glucosami
26	2200	71.2	610	2	Q3EGX5	ACTSC	Q3egx5 actinobacil
27	2183	70.7	609	1	GLMS	PASMU	P57963 p glucosami
28	2166	70.1	610	2	Q47UE2	COLP3	Q47ue2 colwellia p
29	2145	69.4	609	2	Q33SI7	9GAMM	Q33si7 shewanella
30	2145	69.4	609	2	Q3P1V9	9GAMM	Q3plv9 shewanella
31	2142	69.3	609	2	Q3Q819	9GAMM	Q3q819 shewanella

32	2140	69.3	609	2	Q2XIG5	9GAMM	Q2xig5 shewanella
33	2140	69.3	609	2	Q2ZQI1	SHEPU	Q2zqi1 shewanella
34	2135	69.1	609	2	Q368R1	9GAMM	Q368r1 shewanella
35	2134	69.1	609	2	Q2Z4D9	9GAMM	Q2z4d9 shewanella
36	2132	69.0	609	2	Q35V59	9GAMM	Q35v59 shewanella
37	2126	68.8	609	1	GLMS	HAEDU	Q7vkk4 h glucosami
38	2126	68.8	608	1	GLMS	SHEON	Q8cx33 s glucosami
39	2125	68.8	609	2	Q3NKY7	SHEPR	Q3nky7 shewanella
40	2099	68.0	609	1	GLMS	IDILO	Q5qzhs i glucosami
41	2088	67.6	610	2	Q3CSC7	ALTAT	Q3csc7 pseudalter
42	2000	64.8	606	2	Q3IK37	PSEHT	Q3ik37 pseudalter
43	1992	64.5	611	2	Q2XHY7	PSEPU	Q2xhy7 pseudomonas
44	1990	64.4	610	2	Q3K445	PSEPF	Q3k445 pseudomonas
45	1988	64.4	610	1	GLMS	PSEPK	Q88bx8 p glucosami

## ALIGNMENTS

RESULT 1  
Q2M847 ECOLI  
ID Q2M847 ECOLI PRELIMINARY; PRT; 609 AA.  
AC Q2M847;  
DT 21-FEB-2006, integrated into UniProtKB/TrEMBL.  
DT 21-FEB-2006, sequence version 1.  
DT 07-MAR-2006, entry version 2.  
DE L-glutamine:D-fructose-6-phosphate aminotransferase.  
GN Name=glms;  
OS Escherichia coli W3110.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=316407;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=K-12;  
RX MEDLINE=810531692; PubMed=6159575;  
RA Smith D.R., Calvo J.M.;  
RT "Nucleotide sequence of the E coli gene coding for dihydrofolate reductase.";  
RL Nucleic Acids Res. 8:2255-2274(1980).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=K-12;  
RA Sakiguchi T., Ortega-Cesena J., Nosch Y., Ohashi S., Tsuda K., Kanaya S.;  
RT "DNA and amino-acid sequences of 3-isopropylmalate dehydrogenase of Bacillus coagulans. Comparison with the enzymes of Saccharomyces cerevisiae and Thermus thermophilus.";  
RL Biochim. Biophys. Acta 867:36-44(1986).  
RN [3]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=K-12;  
RA Chen H., Sun Y., Stark T., Beattie W., Moses R.E.;  
RT "Nucleotide sequence and deletion analysis of the polB gene of Escherichia coli.";  
RL DNA Cell Biol. 9:613-635(1990).  
RN [4]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=K-12;  
RA Smallshaw J.E., Kelln R.A.;  
RT "Cloning, nucleotide sequence and expression of the Escherichia coli K-12 pyrH gene encoding UMP kinase.";  
RL Genetics (Life Sci. Adv.) 11:59-65(1992).  
RN [5]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=K-12;  
RA Hayaashi K., Morooka N., Yamamoto Y., Fujita K., Isono K., Choi S., Ohtsubo E., Baba T., Wanner B.L., Mori H., Horiuchi T.;  
RT "Highly accurate genome sequences of Escherichia coli K-12 strains MG1655 and W3110.";  
RL Mol. Syst. Biol. 0:0-0(2006).  
RN [6]  
RP NUCLEOTIDE SEQUENCE.

RC STRAIN=K-12;  
 RX PubMed=16397293; DOI=10.1093/nar/gkj150;  
 RA Riley M., Abe T., Arnaud M.B., Berlyn M.K., Blattner F.R.,  
 RA Chaudhuri R.R., Glasner J.D., Horiuchi T., Kessler I.M., Kosuge T.,  
 RA Mori H., Perna N.T., Plunkett G. III, Rudd K.E., Serres M.H.,  
 RA Thomas G.H., Thomson N.R., Wishart D., Wanner B.L.;  
 RT "Escherichia coli K-12: a cooperatively developed annotation snapshot-  
 RT -2005";  
 RL Nucleic Acids Res. 34:1-9 (2006).  
 RN [7]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=K-12;  
 RX MEDLINE=97349980; PubMed=9205837; DOI=10.1093/dnares/4.2.91;  
 RA Yamamoto Y., Aiba H., Baba T., Hayashi K., Inada T., Isono K.,  
 RA Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,  
 RA Mizobuchi K., Mori H., Nakade S., Nakamura Y.,  
 RA Oshima T., Oyama S., Saito N., Sampei G., Satoh Y., Sivasundaram S.,  
 RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,  
 RA Yamagata S., Horiuchi T.;  
 RT "Construction of a contiguous 874-kb sequence of the Escherichia coli  
 RT -K12 genome corresponding to 50.0-68.8 min on the linkage map and  
 RT analysis of its sequence features";  
 RL DNA Res. 4:91-113 (1997).  
 RN [8]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=K-12;  
 RX MEDLINE=97251358; PubMed=9087040; DOI=10.1093/dnares/3.6.379;  
 RA Itoh T., Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,  
 RA Kasai H., Kimura S., Kitakawa M., Kitagawa M., Makino K., Miki T.,  
 RA Mizobuchi K., Mori H., Mori T., Motomura K., Nakade S., Nakamura Y.,  
 RA Nashimoto H., Nishio Y., Oshima T., Saito N., Sampei G., Seki Y.,  
 RA Sivasundaram S., Tagami H., Takeda J., Takemoto K., Wada C.,  
 RA Yamamoto Y., Horiuchi T.;  
 RT "A 460-kb DNA sequence of the Escherichia coli K-12 genome  
 RT corresponding to the 40.1-50.0 min region on the linkage map";  
 RL DNA Res. 3:379-392 (1996).  
 RN [9]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=K-12;  
 RX MEDLINE=97251357; PubMed=9087039; DOI=10.1093/dnares/3.6.363;  
 RA Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,  
 RA Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,  
 RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,  
 RA Nakade S., Nakamura Y., Nashimoto H., Nishio Y., Oshima T., Saito N.,  
 RA Sampei G., Seki Y., Sivasundaram S., Tagami H., Takeda J.,  
 RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;  
 RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome  
 RT corresponding to the 28.0-40.1 min region on the linkage map";  
 RL DNA Res. 3:363-377 (1996).  
 RN [10]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=K-12;  
 RX MEDLINE=97094878; PubMed=8940112; DOI=10.1074/jbc.271.49.31145;  
 RA Arn E.A., Abelson J.N.;  
 RT "The 2'-5' RNA ligase of Escherichia coli. Purification, cloning, and  
 RT genomic disruption";  
 RL J. Biol. Chem. 271:31145-31153 (1996).  
 RN [11]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=K-12;  
 RX MEDLINE=97061202; PubMed=8905232; DOI=10.1093/dnares/3.3.137;  
 RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,  
 RA Ikenoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,  
 RA Kimura S., Kitagawa M., Makino K., Maeda S., Miki T., Mizobuchi K.,  
 RA Mori H., Motomura K., Nakamura Y., Nashimoto H., Nishio Y., Saito N.,  
 RA Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,  
 RA Yano M., Horiuchi T.;  
 RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome  
 RT corresponding to the 12.7-28.0 min region on the linkage map";  
 RL DNA Res. 3:137-155 (1996).  
 RN [12]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=K-12;  
 RX PubMed=94261430; PubMed=8202364;  
 RA Fujita N., Mori H., Yura T., Ishihama A.;  
 RT "Systematic sequencing of the Escherichia coli genome: analysis of the  
 RT 2.4-4.1 min (110,917-193,643 bp) region";  
 RL Nucleic Acids Res. 22:1637-1639 (1994).  
 RN [13]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=K-12;  
 RX MEDLINE=94240115; PubMed=8183897;  
 RA Janosi L., Shimizu I., Kaji A.;  
 RT "Ribosome recycling factor (ribosome releasing factor) is essential  
 RT for bacterial growth";  
 RL Proc. Natl. Acad. Sci. U.S.A. 91:4249-4253 (1994).  
 RN [14]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=K-12;  
 RX MEDLINE=94124004; PubMed=7904973; DOI=10.1016/0378-1119(93)90470-N;  
 RA Allikmets R., Gerrard B.C., Court D., Dean M.C.;  
 RT "Cloning and organization of the abc and mdl genes of Escherichia  
 RT coli: relationship to eukaryotic multidrug resistance";  
 RL Gene 136:231-236 (1993).  
 RN [15]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=K-12;  
 RX MEDLINE=94018640; PubMed=8412694;  
 RA van Heeswijk W.C., Rabenberg M., Westerhoff H.V., Kahn D.D.;  
 RT "The genes of the glutamine synthetase adenylation cascade are not  
 RT regulated by nitrogen in Escherichia coli";  
 RL Mol. Microbiol. 9:443-458 (1993).  
 RN [16]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=K-12;  
 RX MEDLINE=93259920; PubMed=8387990;  
 RA Zhao S., Sandt C.H., Feulner G., Vlazny D.A., Gray J.A., Hill C.W.;  
 RT "Rho elements of Escherichia coli K-12: complex composites of shared  
 RT and unique components that have different evolutionary histories";  
 RL J. Bacteriol. 175:2799-2808 (1993).  
 RN [17]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=K-12;  
 RX MEDLINE=93116053; PubMed=1474579;  
 RA Cormack R.S., Mackie G.A.;  
 RT "Structural requirements for the processing of Escherichia coli 5 S  
 RT ribosomal RNA by RNase E in vitro";  
 RL J. Mol. Biol. 228:1078-1090 (1992).  
 RN [19]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=K-12;  
 RX MEDLINE=93094132; PubMed=1459951;  
 RA Gervais F.G., Drapeau G.R.;  
 RT "Identification, cloning, and characterization of rcsF, a new  
 RT regulator gene for exopolysaccharide synthesis that suppresses the  
 RT division mutation ftsZ84 in Escherichia coli K-12";  
 RL J. Bacteriol. 174:8016-8022 (1992).  
 RN [20]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=K-12;  
 RX MEDLINE=93077430; PubMed=1447125;  
 RA Yamanaka K., Ogura T., Niki H., Hiraga S.;  
 RT "Identification and characterization of the smbA gene, a suppressor of  
 RT the mukB null mutant of Escherichia coli";  
 RL J. Bacteriol. 174:7517-7526 (1992).  
 RN [21]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=K-12;

```
RX MEDLINE=93011013; PubMed=1396599;
RA Condon C., Phillips J., Fu Z.Y., Squires C., Squires C.L.;

Query Match      99.3%; Score 3068; DB 2; Length 609;
Best Local Similarity 99.5%; Pred. No. 7.8e-185;
Matches 606; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MCGIVGAIAORDVAEILLEGRLLEYRGYSAGLAVVDTEGHMTLRRLRGKVMQLAQAAE 60
Db 1 MCGIVGAIAORDVAEILLEGRLLEYRGYSAGLAVVDTEGHMTLRRLRGKVMQLAQAAE 60
Qy 61 EHLPGGTGTAHTRWATHGSPSEVNAHPHYSEHIVVHNGIIEHNEPLRELKARGYTFV 120
Db 61 EHLPGGTGTAHTRWATHGSPSEVNAHPHYSEHIVVHNGIIEHNEPLRELKARGYTFV 120
Qy 121 SETDTEVIAHLVNWELKOGGTLREAVLRAIPQLRGAYGTVMDSRHPDTLAARSGSPLV 180
Db 121 SETDTEVIAHLVNWELKOGGTLREAVLRAIPQLRGAYGTVMDSRHPDTLAARSGSPLV 180
Qy 181 IGLGMGFNTASDQALLPVTRFIFLEEGDIAEITRRSYNIFDKTGAEVKRODIESNLQ 240
Db 181 IGLGMGFNTASDQALLPVTRFIFLEEGDIAEITRRSYNIFDKTGAEVKRODIESNLQ 240
Qy 241 YDAGDKGIYCHMYQKEIYEOPNAIKNTLTGRISHGQVDLSLGNADLLSKVEHIQILA 300
Db 241 YDAGDKGIYCHMYQKEIYEOPNAIKNTLTGRISHGQVDLSLGNADLLSKVEHIQILA 300
Qy 301 CGTSYNGMSRYWFESLAGIPCDVEIASEFRYKSAVRNLSMITLSQSGETADTLAAGL 360
Db 301 CGTSYNGMSRYWFESLAGIPCDVEIASEFRYKSAVRNLSMITLSQSGETADTLAAGL 360
Qy 361 RLSKELGYLSLAICNVPGSSLYRESIALMTNAGTIGVASTKAFQTOLTVLMLVAKL 420
Db 361 RLSKELGYLSLAICNVPGSSLYRESIALMTNAGTIGVASTKAFQTOLTVLMLVAKL 420
Qy 421 SRLKGLDASIEHDIHVHGLQALPSRIEQLSODKRIEALAEFDSKHALFLSRGDQYPIA 480
Db 421 SRLKGLDASIEHDIHVHGLQALPSRIEQLSODKRIEALAEFDSKHALFLSRGDQYPIA 480
Qy 481 LEGALKLKEISYTHAEAYAGELKHGPLALIDAMPVIVVAPNNLEKLSNIEEVRAR 540
Db 481 LEGALKLKEISYTHAEAYAGELKHGPLALIDAMPVIVVAPNNLEKLSNIEEVRAR 540
Qy 541 GGQLYVPADQAGFVSSDNHIIEMPHVEEVIAPFIVTPLLAYHALIKGTVDVQPR 600
Db 541 GGQLYVPADQAGFVSSDNHIIEMPHVEEVIAPFIVTPLLAYHALIKGTVDVQPR 600
Qy 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609

RESULT 2
GLMS_ECOLI
ID _GLMS_ECOLI STANDARD; PRT; 608 AA.
AC P17169; P76745;
DT 01-AUG-1990, integrated into UniProtKB/Swiss-Prot.
DT 15-DEC-1998, sequence version 3.
DT 07-MAR-2006, entry version 70.
DE Glucosamine-6-phosphate aminotransferase (isomerizing)
DE (EC 2.6.1.16) (Hexosephosphate aminotransferase) (D-fructose-6-
DE phosphate amidotransferase) (GFAT) (L-glutamine-D-fructose-6-phosphate
DE amidotransferase) (ordered LocustName=b3729;
GN Name=glms; OrderedGlucosamine=b3729;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
[1]
RN NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RX MEDLINE=85121806; PubMed=6395859;
RA Walker J.E., Gay N.J., Saraste M., Eberle A.N.;
RT "DNA sequence around the Escherichia coli unc operon. Completion of

the sequence of a 17 kilobase segment containing asnA, oriC, unc, glms
and phoS."; 224:799-815 (1984).
[2]
RN NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RC STRAIN=K12 / MG1655;
RX MEDLINE=93315143; PubMed=7686882;
RA Burland V.D., Plunkett G. III, Daniels D.L., Blattner F.R.;
RT "DNA sequence and analysis of 136 kilobases of the Escherichia coli
genome: organizational symmetry around the origin of replication.";
RL Genomics 16:551-561 (1993).
[3]
RN PROTEIN SEQUENCE OF 48-51; 218-230; 488-492; 504-507 AND 600-608.
RX MEDLINE=92007872; PubMed=1915361;
RA Gollinelli-Pimpaneau B., Badet B.;
RT "Possible involvement of lys603 from Escherichia coli glucosamine-6-
RT phosphate synthase in the binding of its substrate fructose 6-
RT phosphate.";
RL Eur. J. Biochem. 201:175-182 (1991).
[4]
RN NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 489-608.
RX MEDLINE=82200022; PubMed=6283361;
RA Lichtenstein C., Bremner S.;
RT "Unique insertion site of Tn7 in the E. coli chromosome.";
RL Nature 297:601-603 (1982).
[5]
RN NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 606-608.
RX MEDLINE=86215091; PubMed=3010949;
RA Gay N.J., Tybuliewicz V.L.J., Walker J.E.;
RT "Insertion of transposon Tn7 into the Escherichia coli glms
RT transcriptional terminator.";
RL Biochem. J. 234:111-117 (1986).
[6]
RN NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 596-608.
RX MEDLINE=88086894; PubMed=2826397;
RA McKown R.L., Orle K.A., Chen T., Craig N.L.;
RT "Sequence requirements of Escherichia coli attTn7, a specific site of
RT transposon Tn7 insertion.";
RL J. Bacteriol. 170:352-358 (1988).
[7]
RN CHARACTERIZATION.
RX MEDLINE=88281539; PubMed=3134953; DOI=10.1016/0300-9084(88)90073-9;
RA Dutka-Mallen S., Mazodier P., Badet B.;
RT "Molecular cloning and overexpression of the glucosamine synthetase
RT gene from Escherichia coli.";
RL Biochimie 70:287-290 (1988).
[8]
RN X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 1-240.
RX MEDLINE=96434326; PubMed=8805567; DOI=10.1016/S0969-2126(96)00087-1;
RA Isupov M.N., Obmolova G., Butterworth S., Badet-Denisot M.-A.,
RA Badet B., Polikarpov I., Littlechild J.A., Teplyakov A.;
RT "Substrate binding is required for assembly of the active conformation
RT of the catalytic site in Ntn amidotransferases: evidence from the 1.8-
RT A crystal structure of the glutaminase domain of glucosamine 6-
RT phosphate synthase.";
RL Structure 4:801-810 (1996).
[9]
RN X-RAY CRYSTALLOGRAPHY (1.57 ANGSTROMS) OF 243-608.
RX MEDLINE=98416699; PubMed=9739095; DOI=10.1016/S0969-2126(98)00105-1;
RA Teplyakov A., Obmolova G., Badet-Denisot M.-A., Badet B.,
RA Polikarpov I.;
RT "Involvement of the C terminus in intramolecular nitrogen channeling
RT in glucosamine 6-phosphate synthase: evidence from a 1.6-A crystal
RT structure of the isomerase domain.";
RL Structure 6:1047-1055 (1998).
[10]
RN X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 241-608.
RX MEDLINE=99190083; PubMed=10091662;
RA Teplyakov A., Obmolova G., Badet-Denisot M.A., Badet B.;
RT "The mechanism of sugar phosphate isomerization by glucosamine 6-
RT phosphate synthase.";
RL Protein Sci. 8:596-602 (1999).
-1- FUNCTION: Catalyzes the first step in hexosamine metabolism,
```

converting fructose-6P into glucosamine-6P using glutamine as a  
 nitrogen source.  
 -!- CATALYTIC ACTIVITY: L-glutamine + D-fructose 6-phosphate = L-  
 glutamate + D-glucosamine 6-phosphate.  
 -!- SUBUNIT: Homodimer.  
 -!- INTERACTION:  
 P0957:al85; NbExp=1; IntAct=EBI-551022, EBI-544061;  
 P08622:dnaj; NbExp=1; IntAct=EBI-551022, EBI-545285;  
 P76552:euht; NbExp=1; IntAct=EBI-551022, EBI-551031;  
 P62615:ispe; NbExp=1; IntAct=EBI-551022, EBI-562202;  
 P61175:rpiv; NbExp=1; IntAct=EBI-551022, EBI-542255;  
 P21166:crkh; NbExp=1; IntAct=EBI-551022, EBI-550268;  
 P76093:ymbd; NbExp=1; IntAct=EBI-551022, EBI-551038;  
 P33366:yohd; NbExp=1; IntAct=EBI-551022, EBI-551046;  
 -!- SUBCELLULAR LOCATION: Cytoplasm.  
 -!- SIMILARITY: In the C-terminal section; belongs to the SIS family.  
 -!- SIMILARITY: In the C-terminal section; belongs to the SIS family.  
 -!- SIMILARITY: Contains 1 type-2 glutamine amidotransferase domain.  
 -----  
 Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 Distributed under the Creative Commons Attribution-NonDerivs License  
 -----  
 EMBL; X01631; CAA25785.1; -; Genomic DNA.  
 EMBL; L10328; AAG62080.1; -; Genomic DNA.  
 EMBL; U00096; AAC76752.1; -; Genomic DNA.  
 EMBL; V00620; CAA23894.1; -; Genomic DNA.  
 EMBL; M18980; AAG23836.1; -; Genomic DNA.  
 PIR; B65176; XNEGCM.  
 DR PDB; 1JXA; X-ray; A/B/C=1-608.  
 DR PDB; 1MOQ; X-ray; @=241-608.  
 DR PDB; 1MOR; X-ray; @=241-608.  
 DR PDB; 1MOS; X-ray; A=241-608.  
 DR PDB; 1XFF; X-ray; A/B=1-240.  
 DR PDB; 1XFG; X-ray; A/B=1-240.  
 DR PDB; 2BPJ; X-ray; A/B=1-608.  
 DR PDB; 2BPL; X-ray; A/B/C=1-608.  
 DR MEROPS; C44.971; -.  
 DR GenomesReviews; U00096\_GR; b3729.  
 DR EcoBASE; EB0377; -.  
 DR EcoGene; EG10382; glms.  
 DR BioCyc; EcoCyc: L-GLN-FRUCT-6-P-AMINOTRANS-MONOMER; -.  
 DR GO; GO:0005515; F:protein binding; IPI.  
 DR HAMAP; MF\_00164; -; 1.  
 DR InterPro; IPR000583; GATase\_2.  
 DR InterPro; IPR005855; Glms\_trans.  
 DR Pfam; PF00310; GATase\_2; 1.  
 DR Pfam; PF01380; SIS; 2.  
 DR TIGRFAMs; TIGR01135; glms; 1.  
 DR PROSITE; P500443; GATASE\_TYPE\_II; 1.  
 3D-structure; Amino transferase; Complete proteome;  
 Direct protein sequencing; Glutamine amidotransferase; Transferase.  
 KW  
 INIT MET 0 0  
 CHAIN 1 608 Glucosamine--fructose-6-phosphate  
 aminotransferase [isomerizing].  
 /FTID-PRO\_0000135328  
 Glutamine amidotransferase.  
 GATase.  
 Isomerization Fru-6P.  
 KL -> NV (in Ref. 1).  
 FT REGION 1 191  
 FT ACT SITE 1 1  
 FT ACT SITE 603 603  
 FT CONFLICT 418 419  
 FT STRAND 2 7  
 FT STRAND 9 10  
 FT STRAND 13 26  
 FT STRAND 29 36  
 FT STRAND 38 39  
 FT STRAND 42 49  
 FT STRAND 51 60  
 FT STRAND 61 61  
 FT STRAND 66 73  
 FT STRAND 76 78  
 FT STRAND 80 81  
 FT STRAND 82 84  
 FT STRAND 85 86

FT STRAND 90  
 FT TURN 91  
 FT STRAND 93  
 FT TURN 101  
 FT TURN 102  
 FT TURN 103  
 FT TURN 104  
 FT TURN 113  
 FT TURN 114  
 FT TURN 115  
 FT STRAND 120  
 FT STRAND 123  
 FT STRAND 123  
 FT STRAND 124  
 FT STRAND 136  
 FT TURN 137  
 FT TURN 137  
 FT STRAND 138  
 FT STRAND 139  
 FT STRAND 141  
 FT STRAND 151  
 FT STRAND 154  
 FT STRAND 162  
 FT TURN 163  
 FT TURN 166  
 FT TURN 167  
 FT STRAND 176  
 FT STRAND 179  
 FT STRAND 184  
 FT STRAND 185  
 FT STRAND 187  
 FT STRAND 192  
 FT STRAND 193  
 FT TURN 196  
 FT TURN 199  
 FT STRAND 201  
 FT STRAND 205  
 FT TURN 208  
 FT STRAND 211  
 FT STRAND 214  
 FT STRAND 216  
 FT STRAND 217  
 FT STRAND 219  
 FT TURN 222  
 FT TURN 224  
 FT STRAND 228  
 FT STRAND 229  
 FT STRAND 233  
 FT STRAND 235  
 FT TURN 246  
 FT STRAND 247  
 FT STRAND 249  
 FT STRAND 250  
 FT STRAND 251  
 FT STRAND 257  
 FT TURN 258  
 FT TURN 259  
 FT TURN 267  
 FT TURN 268  
 FT TURN 269  
 Query Match 99.2%; Score 3063; DB 1; Length 608;  
 Best Local Similarity 99.5%; Pred. No. 1.6e-184; Indels 0; Gaps 0;  
 Matches 605; Conservative 0; Mismatches 3;  
 QY 2 CGIVGAIQRDVAEILLEGRLRLEYRGYDSAGLAVVDTEGHMTRRLRGKVMQLAAAE 61  
 DB 1 CGIVGAIQRDVAEILLEGRLRLEYRGYDSAGLAVVDTEGHMTRRLRGKVMQLAAAE 60  
 QY 62 HPLHGGTGAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHHEPLREELKARGYTFVS 121  
 DB 61 HPLHGGTGAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHHEPLREELKARGYTFVS 120  
 QY 122 ETDTEVIAHLVNWELKQGGTTLREAVLRAIPOLRGAYGTVMDSRHPDTLAARSGSPLVI 181  
 DB 121 ETDTEVIAHLVNWELKQGGTTLREAVLRAIPOLRGAYGTVMDSRHPDTLAARSGSPLVI 180  
 QY 182 GLGMGENFIASDQALLPVTTRFFIFLEEGDIAEITRRSVNIFDKTGAEVKRDIESNLQY 241  
 DB 181 GLGMGENFIASDQALLPVTTRFFIFLEEGDIAEITRRSVNIFDKTGAEVKRDIESNLQY 240  
 QY 242 DAGDKGIYCHVMQKEIYEOPNAIKNTLTGRISHGQVDLSELGPNADLLSKVEHIQILAC 301  
 DB 241 DAGDKGIYCHVMQKEIYEOPNAIKNTLTGRISHGQVDLSELGPNADLLSKVEHIQILAC 300  
 QY 302 GTSYNSGMVSRYPESLAGICPDVEIASEFFRYKSAVRNSLMTLSOSGETADTLAAGLR 361  
 DB 301 GTSYNSGMVSRYPESLAGICPDVEIASEFFRYKSAVRNSLMTLSOSGETADTLAAGLR 360  
 QY 362 LSKELGYLGLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQLTVLLMLVAKLS 421  
 DB 361 LSKELGYLGLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQLTVLLMLVAKLS 420  
 QY 422 RLKGLDASIEHDIHVHGLQALPSRIEOMLSQDKRIEALAEFSDKHALFLSGDQYPIAL 481  
 DB 421 RLKGLDASIEHDIHVHGLQALPSRIEOMLSQDKRIEALAEFSDKHALFLSGDQYPIAL 480  
 QY 482 EGALKLKEISYTHAEAYAAAGELKHGFLALIDADMPVIVVAPNNELLEKLKSNIEEVRAG 541

```

Db 481 EGALKLKEISYTHAEAYAGELKHGPLALIDAMPVIVVAPNNELLEKLSNIEEVRARG 540
Qy 542 GOLYVFADQAGFVSSDNMHIIEMPHVEEVIAPFYTVPLQLLHYHVALIKGTDDVQPRN 601
Db 541 GOLYVFADQAGFVSSDNMHIIEMPHVEEVIAPFYTVPLQLLHYHVALIKGTDDVQPRN 600
Qy 602 LAKSVTVE 609
Db 601 LAKSVTVE 608

RESULT 3
Q329R8_SHIDS
ID Q329R8_SHIDS PRELIMINARY; PRT; 609 AA.
AC Q329R8;
DT 06-DEC-2005, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, entry version 4.
DE L-glutamine:D-fructose-6-phosphate aminotransferase.
GN Name:glms; OrderedLocusNames=SDY_4019;
OS Shigella dysenteriae serotype 1 (strain Sdl97).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=300267;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=16275786; DOI=10.1093/nar/gki954;
RA Yang F., Yang J., Zhang X., Chen L., Jiang Y., Yan Y., Tang X.,
RA Wang J., Xiong Z., Dong J., Xue Y., Zhu Y., Xu X., Sun L., Chen S.,
RA Nie H., Peng J., Xu J., Wang Y., Yuan Z., Wen Y., Yao Z., Shen Y.,
RA Qiang B., Hou Y., Yu J., Jin Q.;
RT "Genome dynamics and diversity of Shigella species, the etiologic
RT agents of bacillary dysentery."
RL Nucleic Acids Res. 33:6445-6458(2005).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; CP000034; ABB63937.1; -; Genomic_DNA.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0004360; F:glutamine-fructose-6-phosphate transaminase. . .; IEA.
DR GO; GO:0005529; F:sugar binding; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0016051; P:carbohydrate biosynthesis; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
KW Amino transferase; Complete proteome; Transferase.
SQ SEQUENCE 609 AA; 65904 MW; 6373301513227484 CRC64;

Query Match 99.2%; Score 3063; DB 2; Length 609;
Best Local Similarity 99.3%; Pred. No. 1.6e-184;
Matches 605; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MCGIVGAIAORDVAEILLEGRLREYRGYDSAGLAVVDTEGHMTRRLRLKQVOMLAQAAE 60
Db 1 MCGIVGAIAORDVAEILLEGRLREYRGYDSAGLAVVDTEGHMTRRLRLKQVOMLAQAAE 60

Qy 61 EPHLHGCTGTAHRWATHGPSEVNAHPHVSEHIVVHNGIIEHPEPLREELKARGYTFV 120
Db 61 EPHLHGCTGTAHRWATHGPSEVNAHPHVSEHIVVHNGIIEHPEPLREELKARGYTFV 120

Qy 121 SETDTEVIAHVNWELKQGTLEAVLRAIPOLRGAYGTVMDSRHPDITLLAARSGSPV 180
Db 121 SETDTEVIAHVNWELKQGTLEAVLRAIPOLRGAYGTVMDSRHPDITLLAARSGSPV 180

Qy 181 IGLGMGNFTASDQALLPVTFRIFLEEGDIAETRRSVNIFDKTGAEVKRODIESNLQ 240
Db 181 IGLGMGNFTASDQALLPVTFRIFLEEGDIAETRRSVNIFDKTGAEVKRODIESNLQ 240

Qy 241 YDAGDKGIYCHYMQEITYEQPNAIKNTLTGRISHGQVDLSLGFNADLLSKVEHIQILA 300
Db 241 YDAGDKGIYCHYMQEITYEQPNAIKNTLTGRISHGQVDLSLGFNADLLSKVEHIQILA 300

```

```

Qy 301 CGTSYNSGMVSRWFESLAGIPCDVEIASEFRYKSAVRENSLMITLSQSGETADTLAGL 360
Db 301 CGTSYNSGMVSRWFESLAGIPCDVEIASEFRYKSAVRENSLMITLSQSGETADTLAGL 360
Qy 361 RLSKELGYLGSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFTTQTLTVLLMLVAKL 420
Db 361 RLSKELGYLGSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFTTQTLTVLLMLVAKL 420
Qy 421 SRLKGLDASIEHDI VHGLQALPSRIEOMLSQDKRIEALAEEDFSDKHHLFLSRGDQYPIA 480
Db 421 SRLKGLDASIEHDI VHGLQALPSRIEOMLSQDKRIEALAEEDFSDKHHLFLSRGDQYPIA 480
Qy 481 LEGALKLKEISYTHAEAYAGELKHGPLALIDAMPVIVVAPNNELLEKLSNIEEVRAR 540
Db 481 LEGALKLKEISYTHAEAYAGELKHGPLALIDAMPVIVVAPNNELLEKLSNIEEVRAR 540
Qy 541 GGQLYVFADQAGFVSSDNMHIIEMPHVEEVIAPFYTVPLQLLHYHVALIKGTDDVQPR 600
Db 541 GGQLYVFADQAGFVSSDNMHIIEMPHVEEVIAPFYTVPLQLLHYHVALIKGTDDVQPR 600
Qy 601 LAKSVTVE 609
Db 601 LAKSVTVE 609

RESULT 4
Q31UM9_SHIDS
ID Q31UM9_SHIDS PRELIMINARY; PRT; 609 AA.
AC Q31UM9;
DT 06-DEC-2005, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, entry version 1.
DE L-glutamine:D-fructose-6-phosphate aminotransferase.
GN Name:glms; OrderedLocusNames=SBO_3758;
OS Shigella boydii serotype 4 (strain Sb227).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=300268;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=16275786; DOI=10.1093/nar/gki954;
RA Yang F., Yang J., Zhang X., Chen L., Jiang Y., Yan Y., Tang X.,
RA Wang J., Xiong Z., Dong J., Xue Y., Zhu Y., Xu X., Sun L., Chen S.,
RA Nie H., Peng J., Xu J., Wang Y., Yuan Z., Wen Y., Yao Z., Shen Y.,
RA Qiang B., Hou Y., Yu J., Jin Q.;
RT "Genome dynamics and diversity of Shigella species, the etiologic
RT agents of bacillary dysentery."
RL Nucleic Acids Res. 33:6445-6458(2005).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; CP000036; ABB68229.1; -; Genomic_DNA.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0004360; F:glutamine-fructose-6-phosphate transaminase. . .; IEA.
DR GO; GO:0005529; F:sugar binding; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0016051; P:carbohydrate biosynthesis; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
KW Amino transferase; Complete proteome; Transferase.
SQ SEQUENCE 609 AA; 66898 MW; 989B94DC79C348D4 CRC64;

Query Match 99.1%; Score 3061; DB 2; Length 609;
Best Local Similarity 99.2%; Pred. No. 2.2e-184;
Matches 604; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MCGIVGAIAORDVAEILLEGRLREYRGYDSAGLAVVDTEGHMTRRLRLKQVOMLAQAAE 60
Db 1 MCGIVGAIAORDVAEILLEGRLREYRGYDSAGLAVVDTEGHMTRRLRLKQVOMLAQAAE 60

Qy 61 EPHLHGCTGTAHRWATHGPSEVNAHPHVSEHIVVHNGIIEHPEPLREELKARGYTFV 120

```

Db 61 EHPHGGTGIAHTRWATHGEPSEANAHPHVSHIMVWHNGIIEHPEUREELKARGYTFV 120  
QY 121 SETDTEVIAHLVNWELKQGTUREAVLRAIPQURGAYGTIVIMDSRHPDTLLAARSGSPVL 180  
Db 121 SETDTEVIAHLVNWELKQGTUREAVLRAIPQURGAYGTIVIMDSRHPDTLLAARSGSPVL 180  
QY 181 IGLGMENFIASDQALLPVTTRFFLEBEGDIAETRRSVNIFDKTGAEVKRDIESNLQ 240  
Db 181 IGLGMENFIASDQALLPVTTRFFLEBEGDIAETRRSVNIFDKTGAEVKRDIESNLQ 240  
QY 241 YDAGDKGIYCHYMOKEIYEPNAINKNTLTGRISHGQVDLSELGPNADLLSKVEHIQILA 300  
Db 241 YDAGDKGIYCHYMOKEIYEPNAINKNTLTGRISHGQVDLSELGPNADLLSKVEHIQILA 300  
QY 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASFEFPRYKSAVRNLSMITLSQSGETADTLA 360  
Db 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASFEFPRYKSAVRNLSMITLSQSGETADTLA 360  
QY 361 RLSKELGYLGSIAICNVPGSSIVRESDLALMTNAGTEIGVASTKAPTQTLTVLLMLVAKL 420  
Db 361 RLSKELGYLGSIAICNVPGSSIVRESDLALMTNAGTEIGVASTKAPTQTLTVLLMLVAKL 420  
QY 421 SRLKGLDASIEHDIYHGLQALPSRIEOMLSQDKRIEALAEFSDKHHLFLSRGDOYPTA 480  
Db 421 SRLKGLDASIEHDIYHGLQALPSRIEOMLSQDKRIEALAEFSDKHHLFLSRGDOYPTA 480  
QY 481 LEGALKLKEISYTHAEVAAAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEEVRAR 540  
Db 481 LEGALKLKEISYTHAEVAAAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEEVRAR 540  
QY 541 GGQLYVFADQDAGFVSSDNNHIIEMPHVEEVIAPITFYVPLQLLAYHVALIKGTVDQPR 600  
Db 541 GGQLYVFADQDAGFVSSDNNHIIEMPHVEEVIAPITFYVPLQLLAYHVALIKGTVDQPR 600  
QY 601 NLAQSVTVE 609  
Db 601 NLAQSVTVE 609  
RESULT 5  
GLMS ECO57  
ID GLMS ECO57 STANDARD; PRT; 608 AA.  
AC Q8XEG2;  
DT 11-JUL-2002, integrated into UniProtKB/Swiss-Prot.  
DT 11-JUL-2002, sequence version 2.  
DT 07-MAR-2006, entry version 24.  
DE Glucosamine-6-phosphate aminotransferase [isomerizing]  
DE (EC 2.6.1.16) (Hexosephosphate aminotransferase) (D-fructose-6-  
DE phosphate amidotransferase) (GFAT) (L-glutamine-D-fructose-6-phosphate  
DE amidotransferase) (Glucosamine-6-phosphate synthase).  
GN Name=glms; OrderedLocustNames=z5227, ECs4671;  
OS Escherichia coli O157:H7.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=83334;  
RN [1]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;  
RX MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;  
RA Perna N.T., Plunkett G., III, Burland V., Mau B., Glasner J.D.,  
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
RA Postai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,  
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
RA Welch R.A., Blattner F.R.;  
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";  
RL Nature 409:529-533 (2001).  
[2]  
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RP STRAIN=O157:H7 / Sakai / RMD 050952 / EHEC;  
RC MEDLINE=21156231; PubMed=11258796; DOI=10.1093/dnares/8.1.11;  
RX Hayaashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,

RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
RA Iida T., Takami H., Honda T., Sasaki C., Ogasawara N., Yasunaga T.,  
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;  
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli  
RT O157:H7 and genomic comparison with a laboratory strain K-12.";  
RL DNA Res. 8:11-22(2001).  
CC -!- FUNCTION: Catalyzes the first step in hexosamine metabolism,  
CC converting fructose-6P into glucosamine-6P using glutamine as a  
CC nitrogen source (By similarity).  
CC -!- CATALYTIC ACTIVITY: L-glutamine + D-fructose 6-phosphate = L-  
CC glutamate + D-glucosamine 6-phosphate.  
CC -!- SUBUNIT: Homodimer (By similarity).  
CC -!- SUBCELLULAR LOCATION: Cytoplasm (By similarity).  
CC -!- SIMILARITY: In the C-terminal section; belongs to the SIS family.  
CC -!- SIMILARITY: Contains 1 type-2 glutamine amidotransferase domain.  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NoDerivs License  
CC  
CC EMBL; AB005174; AAG58932.1; -; Genomic DNA.  
DR EMBL; BA000007; BAB38094.1; -; Genomic DNA.  
DR PIR; G91212; G91212.  
DR PIR; H86058; H86058.  
DR HSP; P17169; 1JXA.  
DR SMR; Q8XEG2; 1-239, 243-608.  
DR GenomeReviews; BA000007 GR; ECs4671.  
DR GenomeReviews; AB005174 GR; z5227.  
DR BiOCYC; ECOL8334-1; ECs4671-MONOMER; -.  
DR HAMAP; MF 00164; -; 1.  
DR InterPro; IPR000583; GATase\_2.  
DR InterPro; IPR005855; GlnS\_trans.  
DR InterPro; IPR001347; SIS.  
DR Pfam; PF00310; GATase\_2; 1.  
DR Pfam; PF01380; SIS; 2.  
DR TIGRFAMs; TIGR01135; glms; 1.  
DR PROSITE; PS00443; GATASE\_TYPE\_II; 1.  
KW Aminotransferase; Complete proteome; Glutamine amidotransferase;  
KW Transferase.  
FT INIT MET 0 0 By similarity.  
FT CHAIN 1 608 Glucosamine--fructose-6-phosphate  
FT aminotransferase [isomerizing].  
FT /FTid=PRO\_0000135330.  
FT REGION 1 240 Glutamine amidotransferase.  
FT ACT\_SITE 1 1 GATase (By similarity).  
FT ACT\_SITE 603 603 Isomerization Fru-6P (By similarity).  
FT CONFLICT 566 566 H -> N (in Ref. 2).  
SQ SEQUENCE 608 AA; 66777 MW; 882EDA38B7F67148 CRC64;  
Query Match 99.0%; Score 3059; DB 1; Length 608;  
Best Local Similarity 99.3%; Pred. No. 2.9e-184;  
Matches 604; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 2 CGIVGAIAQRDVAAEILLEGRLRLEYRGVDSAGLAVVDTEGHMTRRLRLKQVLAQAABE 61  
Db 1 CGIVGAIAQRDVAAEILLEGRLRLEYRGVDSAGLAVVDTEGHMTRRLRLKQVLAQAABE 60  
QY 62 HPLHGGTGIAHTRWATHGEPSEVNAHPHSEHIVVWHNGIIEHPEUREELKARGYTFVS 121  
Db 61 HPLHGGTGIAHTRWATHGEPSEVNAHPHSEHIVVWHNGIIEHPEUREELKARGYTFVS 120  
QY 122 ETDTEVIAHLVNWELKQGTUREAVLRAIPQURGAYGTIVIMDSRHPDTLLAARSGSPVL 181  
Db 121 ETDTEVIAHLVNWELKQGTUREAVLRAIPQURGAYGTIVIMDSRHPDTLLAARSGSPVL 180  
QY 182 IGLGMENFIASDQALLPVTTRFFLEBEGDIAETRRSVNIFDKTGAEVKRDIESNLQ 241  
Db 181 IGLGMENFIASDQALLPVTTRFFLEBEGDIAETRRSVNIFDKTGAEVKRDIESNLQ 240  
QY 242 DAGDKGIYCHYMOKEIYEPNAINKNTLTGRISHGQVDLSELGPNADLLSKVEHIQILAC 301  
Db 241 DAGDKGIYCHYMOKEIYEPNAINKNTLTGRISHGQVDLSELGPNADLLSKVEHIQILAC 300

```

QY 302 GTSYNSGWSRYWPFSLAGIPCDVEIASFPRKSAVRNSLMTTTSQSGETADTLA 361
Db 301 GTSYNSGWSRYWPFSLAGIPCDVEIASFPRKSAVRNSLMTTTSQSGETADTLA 360
QY 362 LSKELGYLGLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFITQTLLV 421
Db 361 LSKELGYLGLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFITQTLLV 420
QY 422 RLKGLDASIEHDIVHGLQALPSRIEQMLSDKRIEALAEAFDSKHHALFLSR 481
Db 421 RLKGLDASIEHDIVHGLQALPSRIEQMLSDKRIEALAEAFDSKHHALFLSR 480
QY 482 EGALKLKEISYTHAEAYAGELKHGPLALIDAMPVIVVAPNNELLEKLKSN 541
Db 481 EGALKLKEISYTHAEAYAGELKHGPLALIDAMPVIVVAPNNELLEKLKSN 540
QY 542 GOLYVFADQAGFVSSDNMHIEMPHVEEVIPIFYTPVQLLAYHVALIKGTD 601
Db 541 GOLYVFADQAGFVSSDNMHIEMPHVEEVIPIFYTPVQLLAYHVALIKGTD 600
QY 602 LAKSVTVE 609
Db 601 LAKSVTVE 608

RESULT 6
GLMS_SHIFL STANDARD; PRT; 608 AA.
AC Q831Y4; Q7BZ99;
DT 30-AUG-2005, integrated into UniProtKB/Swiss-Prot.
DT 30-AUG-2005, sequence version 2.
DT 07-MAR-2006, entry version 22.
DE Glucosamine-6-phosphate aminotransferase (isomerizing)
DE (EC 2.6.1.16) (Hexosephosphate aminotransferase) (D-fructose-6-
DE phosphate aminotransferase) (G6PAT) (L-glutamine-D-fructose-6-phosphate
DE amidotransferase) (Glucosamine-6-phosphate synthase).
GN Names: glms; OrderedLocusNames=SF3809, S3959;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=301 / Serotype 2a; DOI=10.1093/nar/gkf566;
RX MEDLINE=22727406; PubMed=12384590;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157."
RL Nucleic Acids Res. 30:4432-4441(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX MEDLINE=22590274; PubMed=12704152;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.P., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T."
RL Infect. Immun. 71:2775-2786(2003).
CC -I- FUNCTION: Catalyzes the first step in hexosamine metabolism,
CC converting fructose-6P into glucosamine-6P using glutamine as a
CC nitrogen source (By similarity).
CC -I- CATALYTIC ACTIVITY: L-glutamine + D-fructose 6-phosphate = L-
CC glutamate + D-glucosamine 6-phosphate.
CC -I- SUBCELLULAR LOCATION: Cytoplasm (By similarity).
CC -I- SIMILARITY: In the C-terminal section; belongs to the SIS family.
CC G6PAT subfamily.

```

```

CC -I- SIMILARITY: Contains 1 type-2 glutamine amidotransferase domain.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
CC ENBL; AE005674; AAN45249.1; -; Genomic DNA.
DR ENBL; AE016991; AAP18948.1; -; Genomic DNA.
DR HSSP; P17169; 1JXA.
DR SMR; Q831Y4; 1-239, 243-608.
DR GenomeReviews; AE014073_GR; S3959.
DR BioCyc; SFLE198214: AAN45249.1-MONOMER; -.
DR HAMAP; MF 00164; -; 1.
DR InterPro; IPR000583; GATase_2.
DR InterPro; IPR005855; G1ms_trans.
DR Pfam; PF00310; GATase_2; 1.
DR TIGRFAMs; TIGR01135; glms; 1.
DR PROSITE; PS00443; GATASE_TYPE_II; 1.
KW Aminotransferase; Complete proteome; Glutamine amidotransferase;
KW Transferase.
KW INIT MET 0 0 By similarity.
FT CHAIN 1 608 Glucosamine--fructose-6-phosphate
FT FTID-PRO 0000135377,
FT REGION 1 240 Glutamine amidotransferase.
FT ACT_SITE 1 1 GATase (By similarity).
FT ACT_SITE 603 Isomerization Fru-6P (By similarity).
SQ SEQUENCE 608 AA; 66735 MW; B0CEDA38B6F00F7D CRC64;

Query Match 99.0%; Score 3059; DB 1; Length 608;
Best Local Similarity 99.3%; Pred. No. 2.9e-184; Indels 0; Gaps 0;
Matches 604; Conservative 0; Mismatches 4;

QY 2 CGIVGAIAQRDVAEILLEGRLRLEYRGYDSAGLAVVDTEGHMTRRLRLKGVQMLAA 61
Db 1 CGIVGAIAQRDVAEILLEGRLRLEYRGYDSAGLAVVDTEGHMTRRLRLKGVQMLAA 60
QY 62 HPLHGGTGAHTRWATHGEPSEANAPHVSEHIVVHNGIIEHNEPLREELKARGYTFVS 121
Db 61 HPLHGGTGAHTRWATHGEPSEANAPHVSEHIVVHNGIIEHNEPLREELKARGYTFVS 120
QY 122 ETDTEVIHLVNNELKOGGTLREAVLRAIPOLRGAYGTVMDSRHPDTLAARSGSLVI 181
Db 121 ETDTEVIHLVNNELKOGGTLREAVLRAIPOLRGAYGTVMDSRHPDTLAARSGSLVI 180
QY 182 GLGWMGFASDQALLPVTRRFIFLEEGDIABITRRSVNIFDKTGAEVKRQDIENLQY 241
Db 181 GLGWMGFASDQALLPVTRRFIFLEEGDIABITRRSVNIFDKTGAEVKRQDIENLQY 240
QY 242 DAGDKGIYCHYMOKEIVEQNPNAIKNTLTGRISHGQVDLSLGNADLELLSKVHIQILAC 301
Db 241 DAGDKGIYCHYMOKEIVEQNPNAIKNTLTGRISHGQVDLSLGNADLELLSKVHIQILAC 300
QY 302 GTSYNSGWSRYWPFSLAGIPCDVEIASFPRKSAVRNSLMTTTSQSGETADTLA 361
Db 301 GTSYNSGWSRYWPFSLAGIPCDVEIASFPRKSAVRNSLMTTTSQSGETADTLA 360
QY 362 LSKELGYLGLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFITQTLLV 421
Db 361 LSKELGYLGLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFITQTLLV 420
QY 422 RLKGLDASIEHDIVHGLQALPSRIEQMLSDKRIEALAEAFDSKHHALFLSR 481
Db 421 RLKGLDASIEHDIVHGLQALPSRIEQMLSDKRIEALAEAFDSKHHALFLSR 480
QY 482 EGALKLKEISYTHAEAYAGELKHGPLALIDAMPVIVVAPNNELLEKLKSN 541
Db 481 EGALKLKEISYTHAEAYAGELKHGPLALIDAMPVIVVAPNNELLEKLKSN 540
QY 542 GOLYVFADQAGFVSSDNMHIEMPHVEEVIPIFYTPVQLLAYHVALIKGTD 601

```

```

Db      541 GQLYVFADQDAGFVSSDNHIIEMPHVEVIAPIFTYVPLQLLAYHVALIKGTVDVQPRN 600
QY      602 LAKSVTVE 609
      |||||
Db      601 LAKSVTVE 608

RESULT 7
Q3YVN3 SHISS
AC Q3YVN3_SHISS PRELIMINARY; PRT; 609 AA.
DT 27-SEP-2005, integrated into UniProtKB/T-EMBL.
DT 27-FEB-2006, sequence version 1.
DE L-glutamine:D-fructose-6-phosphate aminotransferase.
GN Name=glms; OrderedLocNames=SSO_3890; ORFNames=SSO_3890;
OS Shigella sonnei (strain Ss046).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=300269;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=16275786; DOI=10.1093/nar/gki954;
RA Yang F., Yang J., Zhang X., Chen L., Jiang Y., Yan Y., Tang X.,
RA Wang J., Xiong Z., Dong J., Xue Y., Zhu Y., Xu X., Sun L., Chen S.,
RA Nie H., Peng J., Xu J., Wang Y., Yuan Z., Wen Y., Yao Z., Shen Y.,
RA Qiang B., Hou Y., Yu J., Jin Q.;
RT "Genome dynamics and diversity of Shigella species, the etiologic
RT agents of bacillary dysentery.";
RL Nucleic Acids Res. 33:6445-6458(2005).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CR EMBL; CP000038; AAZ90429.1; -; Genomic_DNA.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0004360; F:glutamine-fructose-6-phosphate transaminase. . .; IEA.
DR GO; GO:0005529; F:sugar binding; IEA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR GO; GO:0016051; P:carbohydrate biosynthesis; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR000593; GATase_2.
DR InterPro; IPR005855; Glms_trans.
DR InterPro; IPR001347; SIS.
DR Pfam; PF00310; GATase_2; 1.
DR Pfam; PF01380; SIS; 2.
DR TIGRFAMs; TIGR01135; glms; 1.
DR PROSITE; PS00443; GATASE_TYPE_II; UNKNOWN 1.
KW Aminotransferase; Complete proteome; Transferase.
SQ SEQUENCE 609 AA; 66844 MW; 9BE75125C4D348D4 CRC64;

Query Match      98.9%; Score 3055; DB 2; Length 609;
Best Local Similarity 99.2%; Pred. No. 5, 1e-184;
Matches 604; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      1 MCGIVGAIAQRDVAIILEGLRLRLEYRGYDSAGLAVVDTGHTMLRLRLGKQVQLAAAE 60
      |||||
Db      1 MCGIVGAIAQRDVAIILEGLRLRLEYRGYDSAGLAVVDAEGHTMLRLRLGKQVQLAAAE 60
      |||||
QY      61 EHPHGGGTGIAHTRWATGPESEVNAHPVSHIVVHNGIIEHPELREELKARGYTFV 120
      |||||
Db      61 EHPHGGGTGIAHTRWATGPESEVNAHPVSHIVVHNGIIEHPELREELKARGYTFV 120
      |||||
QY      121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDPTLLAARSGSLV 180
      |||||
Db      121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDPTLLAARSGSLV 180
      |||||
QY      181 IGLGMGENFIASDQALALPVTRRRFIFLEGDTAEITRRSVNIFDKTGAEVKRDIESNLQ 240
      |||||
Db      181 IGLGMGENFIASDQALALPVTRRRFIFLEGDTAEITRRSVNIFDKTGAEVKRDIESNLQ 240
      |||||
QY      241 YDAGDKGIYRHYMQKEIYEQPNAIKNTLTGRISHGQVDLSELGPNADLLSKVEHIQILA 300
      |||||

```

```

Db      241 YDAGDKGIYRHYMQKEIYEQPNAIKNTLTGRISHGQVDLSELGPNADLLSKVEHIQILA 300
      |||||
QY      301 CGTSYNSGMVSRYPESLAGIPCDVEIASFEFRRKSAVRNLSMITLSQSGETADTLAGL 360
      |||||
Db      301 CGTSYNSGMVSRYPESLAGIPCDVEIASFEFRRKSAVRNLSMITLSQSGETADTLAGL 360
      |||||
QY      361 RLSKELGYLSLAICNVPGSSILVRESDLALMTNAGTEIGVASTKAFPTTQLTVLLMLVAKL 420
      |||||
Db      361 RLSKELGYLSLAICNVPGSSILVRESDLALMTNAGTEIGVASTKAFPTTQLTVLLMLVAKL 420
      |||||
QY      421 SRLKGLDASIEHDIIVHGIQALPSRIEOMLSQDKRIEALAEDEPSDKHHLFLSRGQOYPTA 480
      |||||
Db      421 SRLKGLDASIEHDIIVHGIQALPSRIEOMLSQDKRIEALAEDEPSDKHHLFLSRGQOYPTA 480
      |||||
QY      481 LEGALKLKEISYTHAEAYAGELKHGPLALIDADMPVIVVAPNNLELLEKLSNIEVRRAR 540
      |||||
Db      481 LEGALKLKEISYTHAEAYAGELKHGPLALIDADMPVIVVAPNNLELLEKLSNIEVRRAR 540
      |||||
QY      541 GGQLYVFADQDAGFVSSDNHIIEMPHVEVIAPIFTYVPLQLLAYHVALIKGTVDVQPR 600
      |||||
Db      541 GGQLYVFADQDAGFVSSDNHIIEMPHVEVIAPIFTYVPLQLLAYHVALIKGTVDVQPR 600
      |||||
QY      601 NLAKSVTVE 609
      |||||
Db      601 NLAKSVTVE 609
      |||||

RESULT 8
GLMS_ECOL6
ID GLMS_ECOL6 STANDARD; PRT; 608 AA.
AC Q8FBT4;
DT 06-JUN-2003, integrated into UniProtKB/Swiss-Prot.
DT 06-JUN-2003, sequence version 2.
DT 07-MAR-2006, entry version 17.
DE Glucosamine-fructose-6-phosphate aminotransferase [isomerizing]
DE (EC 2.6.1.16) (Hexosephosphate aminotransferase (D-fructose-6-
DE phosphate amidotransferase) (GfAT) (L-glutamine-D-fructose-6-phosphate
DE amidotransferase) (Glucosamine-6-phosphate synthase).
GN Name=glms; OrderedLocNames=c4654;
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=O6:H1 / CFT073 / ATCC 700928 / UPEC;
RX MEDLINE=22388234; PubMed=12471157; DOI=10.1073/pnas.252529799;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
CC -!- FUNCTION: Catalyzes the first step in hexosamine metabolism,
CC converting fructose-6P into glucosamine-6P using glutamine as a
CC nitrogen source (By similarity).
CC -!- CATALYTIC ACTIVITY: L-glutamine + D-fructose 6-phosphate = L-
CC glutamate + D-glucosamine 6-phosphate.
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasm (By similarity).
CC -!- SIMILARITY: In the C-terminal section; belongs to the SIS family.
CC -!- SIMILARITY: Contains 1 type-2 glutamine amidotransferase domain.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CR EMBL; AE016769; AAN83086.1; -; Genomic_DNA.
DR HSSP; P17169; 1MOS.
DR SMR; Q8FBT4; 1-239, 243-608.
DR GenomeReviews; AE014075_GR; c4654.

```



```
DR BioCyc: ECOL199310:C4654-MONOMER; -
DR HAMAP; MF_00164; -; 1
DR InterPro; IPR000583; GATase_2.
DR InterPro; IPR005855; GIMS_trans.
DR InterPro; IPR001347; SIS.
DR Pfam; PF00310; GATase_2; 1.
DR Pfam; PF01380; GATase_2; 1.
DR TIGRFAMs; TIGR01135; gims; 1.
DR PROSITE; PS00443; GATASE_TYPE_II; 1.
KW Amidotransferase; Complete proteome; Glutamine amidotransferase;
KW Transferase.
FT INIT MET 0 0 By similarity.
FT CHAIN 1 608 Glucosamine--fructose-6-phosphate
FT FT amidotransferase [isomerizing].
FT FT /FTid=PRO_0000135323.
FT REGION 1 240 Glutamine amidotransferase.
FT ACT_SITE 1 1 GATase (By similarity).
FT ACT_SITE 603 603 Isomerization Fru-6P (By similarity).
SQ SEQUENCE 608 AA; 66647 MW; 44DADB5D072C65D7 CRC64;

Query Match 98.7%; Score 3049; DB 1; Length 608;
Best Local Similarity 99.0%; Pred. No. 1.2e-183;
Matches 602; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 2 CGIVGAIQRDVAEILLEGRLRLEYRGYDSAGLAVDTEGHMTLRLRGKVMQLAQAAEE 61
DB 1 CGIVGAIQRDVAEILLEGRLRLEYRGYDSAGLAVDAGHMTLRLRGKVMQLAQAAEE 60
QY 62 HPLHGGTGIATRWATHGEPSEANHPVSEHIVVHNGIENHEPRLKARGYTFVS 121
DB 61 HPLHGGTGIATRWATHGEPSEANHPVSEHIVVHNGIENHEPRLKARGYTFVS 120
QY 122 ETTEVIAHLVNNELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDTLLAARGSPPLVI 181
DB 121 ETTEVIAHLVNNELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDTLLAARGSPPLVI 180
QY 182 GLGWFNFASDQALLPVTRRFFILEEGDIAETTRSVNIFDKTGAEVKRDIESNLQY 241
DB 181 GLGWFNFASDQALLPVTRRFFILEEGDIAETTRSVNIFDKTGAEVKRDIESNLQY 240
QY 242 DAGDKGIYHYMOKIIEYQPNAIKNTLTGRISHGQVDLSBELGPNADLLSKVHEHIQILAC 301
DB 241 DAGDKGIYHYMOKIIEYQPNAIKNTLTGRISHGQVDLSBELGPNADLLSKVHEHIQILAC 300
QY 302 GTSYNSGWSRYWFESLAGIPCDVEIASFRYKSAVRNLSMITLSQSGETADTLAQLR 361
DB 301 GTSYNSGWSRYWFESLAGIPCDVEIASFRYKSAVRNLSMITLSQSGETADTLAQLR 360
QY 362 LSKELGVLGSLAICNVPGSSILRVESDLALMTNAGTEIGVASTKFTTQLTVLLMLVAKLS 421
DB 361 LSKELGVLGSLAICNVPGSSILRVESDLALMTNAGTEIGVASTKFTTQLTVLLMLVAKLS 420
QY 422 RLKGLDASIEHDIHVHGLQALPSRIEQMLSDQKRIEALAEFPDQKHAFSLSRGQDYPAL 481
DB 421 RLKGLDASIEHDIHVHGLQALPSRIEQMLSDQKRIEALAEFPDQKHAFSLSRGQDYPAL 480
QY 482 EGALKLKEISYTHAEVAAAGELKHGPLALIDAMPVIVVAPNNELLKLSNIEEVVARG 541
DB 481 EGALKLKEISYTHAEVAAAGELKHGPLALIDAMPVIVVAPNNELLKLSNIEEVVARG 540
QY 542 GOLYVFADQAGFVSSDNMHIEMPHVEEVTAPFTYVPLQLLAYHVALIKGTDVQDPRN 601
DB 541 GOLYVFADQAGFVSSDNMHIEMPHVEEVTAPFTYVPLQLLAYHVALIKGTDVQDPRN 600
QY 602 LAKSVTVE 609
DB 601 LAKSVTVE 608

RESULT 9
GLMS SALPA STANDARD; PRT; 608 AA.
ID GLMS SALPA
AC QSPK79;
```

```
DT 30-AUG-2005, integrated into UniProtKB/Swiss-Prot.
DT 30-AUG-2005, sequence version 2.
DT 07-MAR-2006, entry version 12.
DE Glucosamine--fructose-6-phosphate aminotransferase [isomerizing]
DE (EC 2.6.1.16) (Hexosephosphate aminotransferase) (D-fructose-6-
DE phosphate amidotransferase) (GFAT) (L-glutamine-D-fructose-6-phosphate
DE amidotransferase) (Glucosamine-6-phosphate synthase).
GN Names:glms; OrderedLocNames=SPA3700;
OS Salmonella paratyphi-a
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=54388;
[1]
NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP STRAIN=ATCC 9150 / SARE42;
RX PubMed=15531882; DOI=10.1038/ng1470;
RA McClelland M., Sanderson K.E., Clifton S.W., Latreille P.,
RA Porwollik S., Sabo A., Meyer R., Bieri T., Ozersky P., McLellan M.,
RA Harkins C.R., Wang C., Nguyen C., Berghoff A., Elliott G.,
RA Kohlberg S., Strong C., Du P., Carter J., Kremizki C., Layman D.,
RA Leonard S., Sun H., Fulton L., Nash W., Miner T., Minx P., Florea L.,
RA Delehaanty K., Fronick C., Magrini V., Nhan M., Warren W., Florea L.,
RA Spieth J., Wilson R.K.;
RT "Comparison of genome degradation in Paratyphi A and Typhi, human-
RL Nat. Genet. 36:1268-1274(2004).
CC -1- FUNCTION: catalyzes the first step in hexosamine metabolism, a
CC . converting fructose-6p into glucosamine-6p using glutamine as a
CC nitrogen source (By similarity).
CC -1- CATALYTIC ACTIVITY: L-glutamine + D-fructose 6-phosphate = L-
CC glutamate + D-glucosamine 6-phosphate.
CC -1- SUBCELLULAR LOCATION: Cytoplasm (By similarity).
CC -1- SIMILARITY: In the C-terminal section; belongs to the SIS family.
CC GFAT subfamily.
CC -1- SIMILARITY: Contains 1 type-2 glutamine amidotransferase domain.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; CP000026; AAV79492.1; -; Genomic_DNA.
CC SMR; QSPKV9; 1-239, 243-608.
CC GenomeReviews; CP000026_GR; SPA3700.
CC HAMAP; MF_00164; -; 1.
CC InterPro; IPR000583; GATase_2.
CC InterPro; IPR005855; GIMS_trans.
CC InterPro; IPR001347; SIS.
CC Pfam; PF00310; GATase_2; 1.
CC Pfam; PF01380; SIS; 2.
CC TIGRFAMs; TIGR01135; gims; 1.
CC PROSITE; PS00443; GATASE_TYPE_II; 1.
KW Amidotransferase; Complete proteome; Glutamine amidotransferase;
KW Transferase.
FT INIT MET 0 0 By similarity.
FT CHAIN 1 608 Glucosamine--fructose-6-phosphate
FT FT amidotransferase [isomerizing].
FT FT /FTid=PRO_0000135373.
FT REGION 1 240 Glutamine amidotransferase.
FT ACT_SITE 1 1 GATase (By similarity).
FT ACT_SITE 603 603 Isomerization Fru-6P (By similarity).
SQ SEQUENCE 608 AA; 66719 MW; BE20C7F1996B742B CRC64;

Query Match 98.4%; Score 3040; DB 1; Length 608;
Best Local Similarity 98.7%; Pred. No. 4.5e-183;
Matches 600; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 2 CGIVGAIQRDVAEILLEGRLRLEYRGYDSAGLAVDTEGHMTLRLRGKVMQLAQAAEE 61
DB 1 CGIVGAIQRDVAEILLEGRLRLEYRGYDSAGLAVDAGHMTLRLRGKVMQLAQAAEE 60
QY 62 HPLHGGTGIATRWATHGEPSEANHPVSEHIVVHNGIENHEPRLKARGYTFVS 121
DB 61 HPLHGGTGIATRWATHGEPSEANHPVSEHIVVHNGIENHEPRLKARGYTFVS 120
```

```
QY 122 ETDTEVIAHLVNWELKQGGTLREAVLRAIPQIRGAYGTVMDSRHPDITLLAARSGSPLVI 181
Db 121 ETDTEVIAHLVNWELKQGGTLREAVLRAIPQIRGAYGTVMDSRHPDITLLAARSGSPLVI 180
QY 182 GLGCMENFIASDQALLPVTTRFFILEEGDIAEITRRSVNIIDKTKGAEVKRDIESNQY 241
Db 181 GLGCMENFIASDQALLPVTTRFFILEEGDIAEITRRSVNIIDKTKGAEVKRDIESNQY 240
QY 242 DAGDKGIYCHYMOKIYEOPNAIKNTLTGRISHGQVDLSLGNADLLEKSKVEHIQILAC 301
Db 241 DAGDKGIYCHYMOKIYEOPNAIKNTLTGRISHGQVDLSLGNADLLEKSKVEHIQILAC 300
QY 302 GTSYNSGMVSRYPFESLAGIPCDVEIASSEFRYKSAVRNLSMITLSQSGTADTLAGLR 361
Db 301 GTSYNSGMVSRYPFESLAGIPCDVEIASSEFRYKSAVRNLSMITLSQSGTADTLAGLR 360
QY 362 LSKELGYLGLSLAICNVPGSSLVRESDLMTNAGTEIGVASTKATTTQTLTVLLMLVAKLS 421
Db 361 LSKELGYLGLSLAICNVPGSSLVRESDLMTNAGTEIGVASTKATTTQTLTVLLMLVAKLS 420
QY 422 RLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEFSDKHAFSLRQDQYPIAL 481
Db 421 RLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEFSDKHAFSLRQDQYPIAL 480
QY 482 EGALKLKEISYTHAEYAAAGELKHGFLALIDADMPIVVVAPNNLEKLSNIEVRARG 541
Db 481 EGALKLKEISYTHAEYAAAGELKHGFLALIDADMPIVVVAPNNLEKLSNIEVRARG 540
QY 542 GQLYVFADQDAGFVSDNNHIIEMPHVEVIAPIFYTVPLQLLAYHVALIKGTDVDDQPN 601
Db 541 GQLYVFADQDAGFVSDNNHIIEMPHVEVIAPIFYTVPLQLLAYHVALIKGTDVDDQPN 600
QY 602 LAKSVTVE 609
Db 601 LAKSVTVE 608

RESULT 10
ID Q57HY2_SALCH PRELIMINARY; PRT; 609 AA.
AC Q57HY2;
DT 10-MAY-2005, integrated into UniProtKB/TrEMBL.
DT 10-MAY-2005, sequence version 1.
DT 07-FEB-2006, entry version 4.
DE L-glutamine:D-fructose-6-phosphate aminotransferase.
GN Name=gms; OrderedLocusNames=SC3774; ORFNames=SCH_3774;
OS Salmonella choleraesuis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=591;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=SC-B67;
RX PubMed=15781495; DOI=10.1093/nar/gki297;
RA Chiu C.-H., Tang P., Chu C., Hu S., Bao Q., Yu J., Chou Y.-Y.,
RA Wang H.-S., Lee Y.-S.;
RT "The genome sequence of Salmonella enterica serovar Choleraesuis, a
RT highly invasive and resistant zoonotic pathogen.";
RL Nucleic Acids Res. 33:1690-1698(2005).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL; AB017220; RA067680.1; -; Genomic_DNA.
DR SMR; Q57HY2; 2-240, 244-609.
DR GO; GO:0005737; C:cycloplasm; IEA.
DR GO; GO:0004360; F:glutamine-fructose-6-phosphate transaminase. . .; IEA.
DR GO; GO:0005529; F:sugar binding; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0016051; P:carbohydrate biosynthesis; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR000583; GATase_2.
```

```
DR InterPro; IPR005855; Gms_trans.
DR InterPro; IPR001347; SIS.
DR Pfam; PF00310; GATase_2; 1.
DR Pfam; PF01380; SIS; 2.
DR TIGRFAMs; TIGR01135; gms; 1.
DR PROSITE; PS00443; GATASE_TYPE_II; UNKNOWN 1.
KW Aminotransferase; Complete proteome; Transferase.
SQ SEQUENCE 609 AA; 66849 MW; EIDIFF6FIAF35F5C CRC64;

Query Match 98.4%; Score 3040; DB 2; Length 609;
Best Local Similarity 98.4%; Pred. No. 4.6e-183;
Matches 599; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 MCGIVGAIQADRVABILLEGRLRREYGYDSAGLAVVDTGHTMLRLRLGKQVLAQAAE 60
Db 1 MCGIVGAIQADRVABILLEGRLRREYGYDSAGLAVVDAEGHMTLRLRLGKQVLAQAAE 60
QY 61 EPHLGGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEENHEPLREELKARGYTFV 120
Db 61 EPHLGGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEENHEPLREELKARGYTFV 120
QY 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPQIRGAYGTVMDSRHPDITLLAARSGSPLV 180
Db 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPQIRGAYGTVMDSRHPDITLLAARSGSPLV 180
QY 181 IGLGMENFIASDQALLPVTTRFFILEEGDIAEITRRSVNIIDKTKGAEVKRDIESNQY 240
Db 181 IGLGMENFIASDQALLPVTTRFFILEEGDIAEITRRSVNIIDKTKGAEVKRDIESNQY 240
QY 241 YDAGDKGIYCHYMOKIYEOPNAIKNTLTGRISHGQVDLSLGNADLLEKSKVEHIQILAC 300
Db 241 YDAGDKGIYCHYMOKIYEOPNAIKNTLTGRISHGQVDLSLGNADLLEKSKVEHIQILAC 300
QY 301 CGTSYNSGMVSRYPFESLAGIPCDVEIASSEFRYKSAVRNLSMITLSQSGTADTLAGL 360
Db 301 CGTSYNSGMVSRYPFESLAGIPCDVEIASSEFRYKSAVRNLSMITLSQSGTADTLAGL 360
QY 361 RLSKELGYLGLSLAICNVPGSSLVRESDLMTNAGTEIGVASTKATTTQTLTVLLMLVAKL 420
Db 361 RLSKELGYLGLSLAICNVPGSSLVRESDLMTNAGTEIGVASTKATTTQTLTVLLMLVAKL 420
QY 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEFSDKHAFSLRQDQYPIA 480
Db 421 ARKLGKLDTSIEHDI VHGLQALPSRIEQMLSDQKRIEALAEFSDKHAFSLRQDQYPIA 480
QY 481 LEGALKLKEISYTHAEYAAAGELKHGFLALIDADMPIVVVAPNNLEKLSNIEVRAR 540
Db 481 LEGALKLKEISYTHAEYAAAGELKHGFLALIDADMPIVVVAPNNLEKLSNIEVRAR 540
QY 541 GGQLYVFADQDAGFVSDNNHIIEMPHVEVIAPIFYTVPLQLLAYHVALIKGTDVDDQPR 600
Db 541 GGQLYVFADQDAGFVSDNNHIIEMPHVEVIAPIFYTVPLQLLAYHVALIKGTDVDDQPR 600
QY 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609

RESULT 11
GLMS_SALTY
ID GLMS_SALTY STANDARD; PRT; 608 AA.
AC Q8ZKX1;
DT 11-JUL-2002, integrated into UniProtKB/Swiss-Prot.
DT 11-JUL-2002, sequence version 2.
DT 07-MAR-2006, entry version 21.
DE Glucosamine-6-phosphate aminotransferase [isomerizing]
DE (EC 2.6.1.16) (Hexosephosphate aminotransferase) (D-fructose-6-
DE phosphate amidotransferase) (GPA) (L-glutamine-D-fructose-6-phosphate
DE amidotransferase) (Glucosamine-6-phosphate synthase).
GN Name=gms; OrderedLocusNames=STM3861;
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
```

```

OX NCBI_TaxID=602;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RA MEDLINE=21534948; PubMed=11677609; DOI=10.1038/35101614;
RA McLeillean M., Sanderson K.E., Spiech J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856(2001).
CC -i- FUNCTION: Catalyzes the first step in hexosamine metabolism,
CC converting fructose-6P into glucosamine-6P using glutamine as a
CC nitrogen source (By similarity).
CC -i- CATALYTIC ACTIVITY: L-glutamine + D-fructose 6-phosphate = L-
CC glutamate + D-glucosamine 6-phosphate.
CC -i- SUBUNIT: Homodimer (By similarity).
CC -i- SUBCELLULAR LOCATION: Cytoplasm (By similarity).
CC -i- SIMILARITY: In the C-terminal section; belongs to the SIS family.
CC -i- SIMILARITY: Contains 1 type-2 glutamine amidotransferase domain.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL: AE008880; AAL22719.1; -; Genomic_DNA.
DR HSSP: P17169; 1JXA.
DR SMR: Q82KX1; 1-239, 243-608.
DR GenomeReviews: AE006468_GR; STM3861.
DR StyGene: SG2222; glms.
DR BioCyc: STY99287:STM3861-MONOMER; -.
DR HAMAP: MF_00164; -; 1.
DR InterPro: IPR000583; GATase_2.
DR InterPro: IPR005855; Glms_trans.
DR InterPro: IPR001347; SIS_trans.
DR Pfam: PF00310; GATase_2; 1.
DR Pfam: PF01380; SIS; 2.
DR TIGRFAMs: TIGR01135; glms; 1.
DR PROSITE: PS00443; GATASE_TYPE_II; 1.
KW Amidotransferase; Complete proteome; Glutamine amidotransferase;
KW Transferase.
FT INIT_MET 0 0 By similarity.
FT CHAIN 1 608 Glucosamine--fructose-6-phosphate
FT FT amidotransferase [isomerizing].
FT FT /FTID=PRO_0000135375.
FT REGION 1 240 Glutamine amidotransferase.
FT ACT_SITE 1 1 GATase (By similarity).
FT ACT_SITE 603 603 Isomerization Fru-6P (By similarity).
SQ SEQUENCE 608 AA; 66746 MW; C320C7F18DEDAF52 CRC64;

Query Match 98.3%; Score 3037; DB 1; Length 608;
Best Local Similarity 98.5%; Pred. No. 7e-183;
Matches 599; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 CGIVGATAQRDVAIBILEGLRLRYRGYDSAGLAVDTEGHTMRLRLGKVMQAQAE 61
DB 1 CGIVGATAQRDVAIBILEGLRLRYRGYDSAGLAVDTEGHTMRLRLGKVMQAQAE 60
QY 62 HPLHGGTGIAHTRWATHGESEVNAHPVSHIIVVHNGIIEHPELREELKARGYTFVS 121
DB 61 HPLHGGTGIAHTRWATHGESEVNAHPVSHIIVVHNGIIEHPELREELKARGYTFVS 120
QY 122 ETDTEVIAHLVNNELKOGGTLREAVLRAIPQLRGAYGTVMDSRHPDTLLAARSGSPLVI 181
DB 121 ETDTEVIAHLVNNELKOGGTLREAVLRAIPQLRGAYGTVMDSRHPDTLLAARSGSPLVI 180
QY 182 GLGNGENFIASDQALLPVTRRPFIFLEGGIAETIRRSVINIFDKTGAEVKRDIESNLQY 241
DB 181 GLGNGENFIASDQALLPVTRRPFIFLEGGIAETIRRSVINIFDKTGAEVKRDIESNLQY 240
QY 242 DAGDKGIYCHYMQKEIYEQNAIKNTLTGRISHGOVDLSLGNADLLSKVEHIQILAC 301

```

```

CC      glutamate + D-glucosamine 6-phosphate.
CC      -!- SUBUNIT: Homodimer (By similarity).
CC      -!- SUBCELLULAR LOCATION: Cytoplasm (By similarity).
CC      -!- SIMILARITY: In the C-terminal section; belongs to the SIS family.
CC      GFAT subfamily.
CC      -!- SIMILARITY: Contains 1 type-2 glutamine amidotransferase domain.
CC      -----
CC      Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC      Distributed under the Creative Commons Attribution-NoDerivs License
CC      -----
DR      EMBL; AL627280; CAD03134.1; -; Genomic DNA.
DR      EMBL; AE014613; AAO71155.1; -; Genomic_DNA.
DR      HSSP; P17169; 1MOS.
DR      Q82ZQ2; 1-239, 243-608.
DR      GenomeReviews; AL513382.GR; STY3917.
DR      GenomeReviews; AE014613.GR; T3658.
DR      BioCVC; SENT209261.T3658-MONOMER; -.
DR      BioCVC; SENT90370.STY3917-MONOMER; -.
DR      HAMAP; MF_00164; -; 1.
DR      InterPro; IPR000583; GATase_2.
DR      InterPro; IPR005855; Glns_trans.
DR      InterPro; IPR001347; SIS.
DR      Pfam; PF00310; GATase_2; 1.
DR      Pfam; PF01380; SIS; 2.
DR      Pfam; PF01380; SIS; 2.
DR      TIGRFAMs; TIGR01135; glns; 1.
DR      PROSITE; PS00443; GATASE_TYPE_II; 1.
DR      AminoTransferase; Complete proteome; Glutamine amidotransferase;
KW      Transferase.
FT      INIT MET      0      0
FT      CHAIN         1      608
FT
FT      REGION      1      240
FT      ACT SITE    1      1
FT      ACT SITE    603   603
FT      ACT SITE    608   608
FT      SEQUENCE    608   AA; 66721 MW; 3F18E4BA47D4A804 CRC64;
SQ
Query Match      98.1%; Score 3030; DB 1; Length 608;
Best Local Similarity 98.0%; Pred. No. 1.9e-182;
Matches 596; Conservative 5; Mismatches 7; Indels 0; Gaps 0;
QY      2  CGIVGAIAQDVAEILLEGRLRYGVDSAGLVVDTGHTRLRLRGKVMQLAAQAE 61
DB      1  CGIVGAIAQDVAEILLEGRLRYGVDSAGLVVDTGHTRLRLRGKVMQLAAQAE 60
QY      62  HPLHGGTGIATRWATHGSPSEVNAHPVSEHIVVYVHNGIIEHPELRKARGYTFVS 121
DB      61  HPLHGGTGIATRWATHGSPSEANAHVPVSEHIVVYVHNGIIEHPELRKARGYTFVS 120
QY      122  ETDTEVIAHLVNWELKQGTLEAVLRAIPQLRGAYGTVIMDSRPHDTLLAARSGSPLVI 181
DB      121  ETDTEVIAHLVNWELKQGTLEAVLRAIPQLRGAYGTVIMDSRPHDTLLAARSGSPLVI 180
QY      182  GLGMGNFTASQDALLPVTREFFILEEGDIAETTRSVNIFDKTCAEVKRODIESNLOY 241
DB      181  GLGMGNFTASQDALLPVTREFFILEEGDIAETTRSVNIFDKTCAEVKRODIESNLOY 240
QY      242  DAGDGIYCHYNQKEIYEOPNAIKNTLTGRIISHGOVDLSSELGNADLLSKVEHIQILAC 301
DB      241  DAGDGIYCHYNQKEIYEOPNAIKNTLTGRIISHGOVDLSSELGNADLLSKVEHIQILAC 300
QY      302  GTSYNSGMVSRVWFESLAGICDVEIASEFRYKSAVRNSLMITLSQSGETADTLAIGLR 361
DB      301  GTSYNSGMVSRVWFESLAGICDVEIASEFRYKSAVRNSLMITLSQSGETADTLAIGLR 360
QY      362  LSKELGYLGSIAICNVPGSSLVRESDLMTNAGTEIGVASTKFTTQTLTLLMLVAKLS 421
DB      361  LSKELGYLGSIAICNVPGSSLVRESDLMTNAGTEIGVASTKFTTQTLTLLMLVAKLS 420
QY      422  RLKGLDASTIEHDIVHGLQALPSRIEQLMSQDKRIEALAEFDSKHALFLSRGDQYPIAL 481
DB      421  RLKGLDASTIEHDIVHGLQALPSRIEQLMSQDKRIEALAEFDSKHALFLSRGDQYPIAL 480

```

```

QY      482  EGALKLKEISYTHAEYAAAGELKHGKPLALIDADMPVIVVAPNNELEKLSNIEVRARG 541
DB      481  EGALKLKEISYTHAEYAAAGELKHGKPLALIDADMPVIVVAPNNELEKLSNIEVRARG 540
QY      542  GOLYVFADQDAGFVSSDDNNHIIEMPHVEVIAPIFYTVPLQLLAYHVALIKGTDVDPNRN 601
DB      541  GOLYVFADQDAGFVSSDDNNHIIEMPHVEVIAPIFYTVPLQLLAYHVALIKGTDVDPNRN 600
QY      602  LAKSVTVE 609
DB      601  LAKSVTVE 608
RESULT 13
GLMS_ERWCT
ID      GLMS_ERWCT      STANDARD;      PRT;      609   AA.
AC      Q6CVJ9;
DT      30-AUG-2005, integrated into UniProtKB/Swiss-Prot.
DT      30-AUG-2005, sequence version 2.
DE      07-MAR-2006, entry version 15.
DE      Glucosamine--fructose-6-phosphate aminotransferase [isomerizing]
DE      (EC 2.6.1.16) (Hexosephosphate aminotransferase) (D-fructose-6-
DE      phosphate amidotransferase) (GFAT) (L-glutamine-D-fructose-6-phosphate
DE      amidotransferase) (Glucosamine-6-phosphate synthase).
DE      Name=glns; OrderedLocustNames=ECA4508;
OS      Erwinia carotovora subsp. atroseptica (Pectobacterium atrosepticum).
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC      Enterobacteriaceae; Pectobacterium.
OX      NCBI_TaxID=29471;
RN      [1]_
RP      NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC      STRAIN=SCRI 1043 / ATCC BAA-672;
RX      PubMed=15263089; DOI=10.1073/pnas.0402424101;
RA      Bell K.S., Sebailia M., Pritchard L., Holden M.T.G., Hyman L.J.,
RA      Holeva M.C., Thomson N.R., Bentley S.D., Churcher L.J.C., Mungall K.,
RA      Atkin R., Bason N., Brooks K., Chillingworth T., Clark K., Doggett J.,
RA      Frazer A., Hance Z., Hauser H., Jagsels K., Moule S., Norbertozak H.,
RA      Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,
RA      Salmond G.P.C., Birch P.R.J., Parkhill J., Toth I.K.;
RT      "Genome sequence of the enterobacterial phytopathogen Erwinia
RT      carotovora subsp. atroseptica and characterization of virulence
RT      factors.";
RL      Proc. Natl. Acad. Sci. U.S.A. 101:11105-11110(2004).
CC      -!- FUNCTION: Catalyzes the first step in hexosamine metabolism,
CC      converting fructose-6p into glucosamine-6p using glutamine as a
CC      nitrogen source (By similarity).
CC      -!- CATALYTIC ACTIVITY: L-glutamine + D-fructose 6-phosphate = L-
CC      glutamate + D-glucosamine 6-phosphate.
CC      -!- SUBCELLULAR LOCATION: Cytoplasm (By similarity).
CC      -!- SIMILARITY: In the C-terminal section; belongs to the SIS family.
CC      GFAT subfamily.
CC      -!- SIMILARITY: Contains 1 type-2 glutamine amidotransferase domain.
CC      -----
CC      Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC      Distributed under the Creative Commons Attribution-NoDerivs License
CC      -----
DR      EMBL; BX950851; CAG77403.1; -; Genomic_DNA.
DR      SMR; Q6CVJ9; 1-239, 244-609.
DR      GenomeReviews; BX950851.GR; ECA4508.
DR      HAMAP; MF_00164; -; 1.
DR      InterPro; IPR000583; GATase_2.
DR      InterPro; IPR005855; Glns_trans.
DR      InterPro; IPR001347; SIS.
DR      Pfam; PF00310; GATase_2; 1.
DR      Pfam; PF01380; SIS; 2.
DR      TIGRFAMs; TIGR01135; glns; 1.
DR      PROSITE; PS00443; GATASE_TYPE_II; 1.
KW      AminoTransferase; Complete proteome; Glutamine amidotransferase;
KW      Transferase.
FT      INIT MET      0      0
FT      CHAIN         1      609
FT
FT      REGION      1      609
FT      ACT SITE    1      1
FT      ACT SITE    609   609
FT      ACT SITE    609   609
FT      SEQUENCE    609   AA; 66721 MW; 3F18E4BA47D4A804 CRC64;
SQ
Query Match      98.1%; Score 3030; DB 1; Length 609;
Best Local Similarity 98.0%; Pred. No. 1.9e-182;
Matches 596; Conservative 5; Mismatches 7; Indels 0; Gaps 0;
QY      2  CGIVGAIAQDVAEILLEGRLRYGVDSAGLVVDTGHTRLRLRGKVMQLAAQAE 61
DB      1  CGIVGAIAQDVAEILLEGRLRYGVDSAGLVVDTGHTRLRLRGKVMQLAAQAE 60
QY      62  HPLHGGTGIATRWATHGSPSEVNAHPVSEHIVVYVHNGIIEHPELRKARGYTFVS 121
DB      61  HPLHGGTGIATRWATHGSPSEANAHVPVSEHIVVYVHNGIIEHPELRKARGYTFVS 120
QY      122  ETDTEVIAHLVNWELKQGTLEAVLRAIPQLRGAYGTVIMDSRPHDTLLAARSGSPLVI 181
DB      121  ETDTEVIAHLVNWELKQGTLEAVLRAIPQLRGAYGTVIMDSRPHDTLLAARSGSPLVI 180
QY      182  GLGMGNFTASQDALLPVTREFFILEEGDIAETTRSVNIFDKTCAEVKRODIESNLOY 241
DB      181  GLGMGNFTASQDALLPVTREFFILEEGDIAETTRSVNIFDKTCAEVKRODIESNLOY 240
QY      242  DAGDGIYCHYNQKEIYEOPNAIKNTLTGRIISHGOVDLSSELGNADLLSKVEHIQILAC 301
DB      241  DAGDGIYCHYNQKEIYEOPNAIKNTLTGRIISHGOVDLSSELGNADLLSKVEHIQILAC 300
QY      302  GTSYNSGMVSRVWFESLAGICDVEIASEFRYKSAVRNSLMITLSQSGETADTLAIGLR 361
DB      301  GTSYNSGMVSRVWFESLAGICDVEIASEFRYKSAVRNSLMITLSQSGETADTLAIGLR 360
QY      362  LSKELGYLGSIAICNVPGSSLVRESDLMTNAGTEIGVASTKFTTQTLTLLMLVAKLS 421
DB      361  LSKELGYLGSIAICNVPGSSLVRESDLMTNAGTEIGVASTKFTTQTLTLLMLVAKLS 420
QY      422  RLKGLDASTIEHDIVHGLQALPSRIEQLMSQDKRIEALAEFDSKHALFLSRGDQYPIAL 481
DB      421  RLKGLDASTIEHDIVHGLQALPSRIEQLMSQDKRIEALAEFDSKHALFLSRGDQYPIAL 480

```

```
FT REGION 1 241 Glutamine amidotransferase.
FT ACT_SITE 1 1 GATase (By similarity).
FT ACT_SITE 604 604 Isomerization Fru-6P (By similarity).
SQ SEQUENCE 609 AA; 6686 MW; BFE71B8EDFD23B5D CRC64;

Query Match
Best Local Similarity 86.4%; Score 2667.5; DB 1; Length 609;
Matches 516; Conservative 46; Mismatches 46; Indels 1; Gaps 1;

QY 2 CGIVGAIQADVAIEILLEGRLRLRYGYDSAGLAVDTEGHMTLRRLGKVKQMLAQAAEE 61
Db 1 CGIVGAVAQDVAIEILLEGRLRLRYGYDSAGLAVDSEGHVARLRRLGKVKQVLSQAAEE 60
QY 62 HPLHGGTGIAHTRWATGCESEVNAHPVSEHIVVWNGIIEHNEPRLBELKRGYTFVS 121
Db 61 HELHGGTGIAHTRWATGCESEENAPHISEHTITVHNGIIEHNEPRLBELMIGRGYFVS 120
QY 122 ETDTEVIAHLVNNELKQ-CGTLREAVLRAIPQLRGAYGTVMDSRHPDTLLAARSGSPLV 180
Db 121 ETDTEVIAHLVHPEQKQNGTGVVVKRVIPQLRGAYGMVLDNRDSSVLVAARSGSPLV 180
QY 181 IGLGMGNFIASDQALLPVTRPFIPEEGDIABITRRSVNIIDTKGAEVKRODIESNLQ 240
Db 181 IGRGVGNFIASDQALLPVTRPFIPEEGDVAEITRRDVRVFDKSGQLATREEIESKVS 240
QY 241 YDAGDKIYCHYQKEIYEQNAIKNTLTGRI SHGOVDLSLGNADNELLKVEHIOILA 300
Db 241 YDAGDKGAYRHYQKEIYEQPMAIKNTLTGRI SHGFIINLSELGPKADNELLKVEHVQIIA 300
QY 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASFRYKRSVNRNSLMTLSQSGETADTLAAGL 360
Db 301 CGTSYNSGMVSRVWFELAGIPCDVEIASFRYKRPVNRNSLMTLSQSGETADTLAAL 360
QY 361 RLSELGYLGLSLAICNVPGSSLVRESPLALMTNAGTEIGVASTKAFQTQTLVLLMLVAKL 420
Db 361 RLSELGYLGLSLAICNVAGSSLVRESPLALMTNAGTEIGVASTKAFQTQTLVLLMLVARV 420
QY 421 SRLKGLDASTEHDIHVHGLQALPSRIEQLMSQDKRIEALAEFDSFKHAFILSRGDQYPIA 480
Db 421 GRUGMDAQTEHDIHVHGLQALPARIEQLMSQDKLIESAEFGFSFKHAFILGRGDQYPIA 480
QY 481 LEGALKLKEISYTHAEAYAGELKHGFLALIDAMPVIVVAPNNLEKLKSNIEEVRAR 540
Db 481 MEGALKLKEISYTHAEAYAGELKHGFLALIDAMPVIVVAPNNLEKLKSNIEEVRAR 540
QY 541 GGQLYVFADQDQGVSSDNNHIIEMPHVEVIAPIFYTVPLQLLAYHVALIKGTDVDQPR 600
Db 541 GGELYVFADEDAGTSSNNKIIPLPHIEVIAPIFYTVPLQLLHYHVALIKGTDVDQPR 600
QY 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609

RESULT 14
GLMS_YERPE STANDARD; PRT; 608 AA.
AC Q829S8;
DT 11-JUL-2002, integrated into UniProtKB/Swiss-Prot.
DT 11-JUL-2002, sequence version 2.
DT 07-MAR-2006, entry version 32.
DE Glucosamine-6-phosphate aminotransferase [isomerizing]
DE (EC 2.6.1.16) (Hexosephosphate aminotransferase) (D-fructose-6-
DE phosphate amidotransferase) (GFPAT) (L-glutamine-D-fructose-6-phosphate
DE amidotransferase) (Glucosamine-6-phosphate synthase).
GN Name:glms; OrderedLocusNames=YPO4118, y4132, YP4025;
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP STRAIN=CO-92 / Biovar Orientalis;
RC
```

```

RX MEDLINE=21470413; PubMed=11586360; DOI=10.1038/35097083;
RA Parkhill J., Wren B.W., Thomson N.R., Tittball R.W., Holden M.T.G.,
RA Prentice M.B., Sebaihia M., James K.D., Churcher C.M., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.-M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.M.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrall B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague."
RL Nature 413:523-527(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=KIM5 / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RX DOI=10.1128/JB.184.16.4601-4611.2002;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM."
RL J. Bacteriol. 184:4601-4611(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=91001 / Biovar Mediaevalis;
RX PubMed=15368893; DOI=10.1093/dnares/11.3.179;
RA Song Y., Tong Z., Wang J., Wang L., Guo Z., Han Y., Zhang J., Pei D.,
RA Zhou D., Qin H., Pang X., Han Y., Zhai J., Li M., Cui B., Qi Z.,
RA Jin L., Dai R., Chen F., Li S., Ye C., Du Z., Lin W., Wang J., Yu J.,
RA Yang H., Wang J., Huang P., Yang R.;
RT "Complete genome sequence of Yersinia pestis strain 91001, an isolate
RT avirulent to humans."
RL DNA Res. 11:179-197(2004).
CC -1- FUNCTION: Catalyzes the first step in hexosamine metabolism,
CC converting fructose-6p into glucosamine-6p using glutamine as a
CC nitrogen source (By similarity).
CC -1- CATALYTIC ACTIVITY: L-glutamine + D-fructose 6-phosphate = L-
CC glutamate + D-glucosamine 6-phosphate.
CC -1- SUBCELLULAR LOCATION: Cytoplasm (By similarity).
CC -1- SIMILARITY: In the C-terminal section; belongs to the SIS family.
CC GFPAT subfamily.
CC -1- SIMILARITY: Contains 1 type-2 glutamine amidotransferase domain.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonCommercial License
CC
CC -----
CC EMBL: AJ414160; CAC93567.1; -; Genomic DNA.
CC EMBL: AE014014; AAM87674.1; -; Genomic DNA.
CC EMBL: AE017142; AAS64164.1; -; Genomic DNA.
CC PIR: AB0500; AB0500.
CC HSP: P17169; IMOS.
CC SNR: Q829S8; 1-239, 243-608.
CC GenomeReviews: AE009952 GR; Y4132.
CC GenomeReviews: AE017042 GR; YP4025.
CC GenomeReviews: AL590842 GR; YPO4118.
CC BioCyc: YPES187410.Y4132-MONOMER; -.
CC BioCyc: YPES632:YPO4118-MONOMER; -.
CC HAMAP: MF_00164; -.
CC InterPro: IPR000583; GATase_2.
CC InterPro: IPR005855; Glms_trans.
CC InterPro: IPR001347; SIS_trans.
CC Pfam: PF00310; GATase_2; 1.
CC Pfam: PF01380; SIS; 2.
CC TIGRFAMs: TIGR01135; glms; 1.
CC PROSITE: PS00443; GATASE_TYPE_II; 1.
CC AminoTransfrase; Complete proteome; Glutamine amidotransferase;
KW Transfrase.
FT INIT_MET 0 0 By similarity.
FT CHAIN 1 608 Glucosamine--fructose-6-phosphate
FT aminotransferase [isomerizing].
FT /FTID=PRO_0000135418.
FT REGION 1 240 Glutamine amidotransferase.
FT ACT_SITE 1 1 GATase (By similarity).
FT
```

```
FT ACT SITE 603 603 Isomerization Fru-6P (By similarity).
SQ SEQUENCE 608 AA; 66394 MW; 1655FD0ADB16CCD6 CRC64;

Query Match
Best Local Similarity 86.2%; Score 2663; DB 1; Length 608;
Matches 515; Conservative 44; Mismatches 49; Indels 0; Gaps 0;

QY 2 CGIVGAIAQRDVAAEILLEGRLRLEVRGYDSAGLAVVDTEGHTRRLRRGKVMQLAAAE 61
Db 1 CGIVGAIAQRDIAEILLEGRLRLEVRGYDSAGLAVVDSEGLTRLRVRGKVALSDAAEK 60
QY 62 HPLHGGTGAIAHTRWATHGEPSEVNAHPVSEHVVVHNGIIEHHEPLREELKARGYTFVS 121
Db 61 QDLHGGTGAIAHTRWATHGEPSEANAHPHVSDYSVVHNGIIEHHEPLRELLISRGYRFS 120
QY 122 ETDTEVIAHLVNWELKQGTLEAREVLRAIPOLRGAYGTVMDSRHPDITLLAARSGSPVLI 181
Db 121 ETDTEVIAHLVHWEQOGGSLLEVVKRVIPOLRGAYGTVMDSRDPSPRLIAARSGSPVLI 180
QY 182 GLGMGENFIASDQALLPVTFRFIFLEEGDIAETRRSVNIPDKTGAEVKRODIESNLOY 241
Db 181 GCGVGENFIASDQALLPVTFRFIFLEEGDVVEVTRRSISIFDKQGNAIERPEIESOVQY 240
QY 242 DAGDKGIYCHYMQEIEYQPNIAKNTLTGRISHGOVDLSELGPNADLLSKVEHIQILAC 301
Db 241 DAGDKGIYRHYMQEIEYQPNIAKNTLTGRLSHGMDLSELGPKADALLAEVQHIQIIAC 300
QY 302 GTSYNSGMVSRVYWFESLAGIPCDVEIASFEFRYKSAVRNSLMTLSQSGETADTLAAGR 361
Db 301 GTSYNSGMVSRVYWFESLAGVPCDVEIASFEFRYKSAVRNSLMTLSQSGETADTLAALR 360
QY 362 LSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQTLTVLLMLVAKUS 421
Db 361 LSKELGYLGSIAICNVAGSSLVRESDLALMTKAGTEIGVASTKFTTQTLTVLLMLVGRIG 420

RESULT 15
GLMS_YERPS
ID GLMS_YERPS STANDARD; PRT; 608 AA.
AC Q663R1;
DT 30-AUG-2005, integrated into UniProtKB/Swiss-Prot.
DT 30-AUG-2005, sequence version 2.
DE Glucosamine-6-phosphate aminotransferase [isomerizing]
DE (EC 2.6.1.16) (Hexosephosphate aminotransferase) (D-fructose-6-
DE phosphate amidotransferase) (G6P) (L-glutamine-D-fructose-6-phosphate
DE amidotransferase) (Glucosamine-6-phosphate synthase).
GN Name=g1ms; OrderedLocustNames=yPTB3964;
OS Yersinia pseudotuberculosis.
OC Bacteria; Proteobacteria; Gammaaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=633;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=IP32953 / Serotype I;
RX PubMed=15358858; DOI=10.1073/pnas.0404012101;
RA Chain P.S.G., Carniel E., Larimer P.W., Lamerdin J., Stoutland P.O.,
```

```
RA Regala W.M., Georgescu A.M., Vergez L.M., Land M.L., Motin V.L.,
RA Brubaker R.R., Fowler J., Hinnebusch J., Marceau M., Medigue C.,
RA Simonet M., Chenal-Francois V., Souza B., Dacheux D., Elliott J.M.,
RA Derbise A., Hauser L.J., Garcia E.;
RT "Insights into the evolution of Yersinia pestis through whole-genome
RT comparison with Yersinia pseudotuberculosis.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:13826-13831(2004).
CC -!- FUNCTION: Catalyzes the first step in hexosamine metabolism,
CC converting fructose-6P into glucosamine-6P using glutamine as a
CC nitrogen source (By similarity).
CC -!- CATALYTIC ACTIVITY: L-glutamine + D-fructose 6-phosphate = L-
CC glutamate + D-glucosamine 6-phosphate.
CC -!- SUBCELLULAR LOCATION: Cytoplasm (By similarity).
CC -!- SIMILARITY: In the C-terminal section; belongs to the SIS family.
CC GFAT subfamily.
CC -!- SIMILARITY: Contains 1 type-2 glutamine amidotransferase domain.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; BX936398; CAH23202.1; -; Genomic_DNA.
CC SMR; Q663R1; 1-239, 243-608.
CC GenomeReviews; BX936398 GR; YPTB3964.
CC HAMAP; MF_00164; -; 1.
CC InterPro; IPR000583; GATase_2.
CC InterPro; IPR005855; G1ms_trans.
CC InterPro; IPR001347; SIS.
CC Pfam; PF00310; GATase_2; 1.
CC TIGRFAMs; TIGR01135; g1ms; 1.
CC PROSITE; PS00443; GATASE_TYPE_II; 1.
KW Aminotransferase; Complete proteome; Glutamine amidotransferase;
KW Transferase.
FT INIT MET 0 0 By similarity.
FT CHAIN 1 608 Glucosamine--fructose-6-phosphate
FT aminotransferase [isomerizing].
FT /FTId=PRO_0000135419.
FT REGION 1 240 Glutamine amidotransferase.
FT ACT SITE 1 1 GATase (By similarity).
FT ACT SITE 603 603 Isomerization Fru-6P (By similarity).
SQ SEQUENCE 608 AA; 66394 MW; 1655FD0ADB16CCD6 CRC64;

Query Match
Best Local Similarity 86.2%; Score 2663; DB 1; Length 608;
Matches 515; Conservative 44; Mismatches 49; Indels 0; Gaps 0;

QY 2 CGIVGAIAQRDVAAEILLEGRLRLEVRGYDSAGLAVVDTEGHTRRLRRGKVMQLAAAE 61
Db 1 CGIVGAIAQRDIAEILLEGRLRLEVRGYDSAGLAVVDSEGLTRLRVRGKVALSDAAEK 60
QY 62 HPLHGGTGAIAHTRWATHGEPSEVNAHPVSEHVVVHNGIIEHHEPLREELKARGYTFVS 121
Db 61 QDLHGGTGAIAHTRWATHGEPSEANAHPHVSDYSVVHNGIIEHHEPLRELLISRGYRFS 120
QY 122 ETDTEVIAHLVNWELKQGTLEAREVLRAIPOLRGAYGTVMDSRHPDITLLAARSGSPVLI 181
Db 121 ETDTEVIAHLVHWEQOGGSLLEVVKRVIPOLRGAYGTVMDSRDPSPRLIAARSGSPVLI 180
QY 182 GLGMGENFIASDQALLPVTFRFIFLEEGDIAETRRSVNIPDKTGAEVKRODIESNLOY 241
Db 181 GCGVGENFIASDQALLPVTFRFIFLEEGDVVEVTRRSISIFDKQGNAIERPEIESOVQY 240
QY 242 DAGDKGIYCHYMQEIEYQPNIAKNTLTGRISHGOVDLSELGPNADLLSKVEHIQILAC 301
Db 241 DAGDKGIYRHYMQEIEYQPNIAKNTLTGRLSHGMDLSELGPKADALLAEVQHIQIIAC 300
QY 302 GTSYNSGMVSRVYWFESLAGIPCDVEIASFEFRYKSAVRNSLMTLSQSGETADTLAAGR 361
Db 301 GTSYNSGMVSRVYWFESLAGVPCDVEIASFEFRYKSAVRNSLMTLSQSGETADTLAALR 360
QY 362 LSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQTLTVLLMLVAKUS 421
Db 361 LSKELGYLGSIAICNVAGSSLVRESDLALMTKAGTEIGVASTKFTTQTLTVLLMLVGRIG 420
```

Qy	422	RLKGLDASIEHDI VHGI QALPSRIEOMLSODKRIEALAEDEFS DKHHALFLSRGDQYPIAL	481
Db	421	KLKGADASLEHDI VHALQALPARIEQMLSLDKTIEALAEGFS DKHHALFLGRGDQYPIAM	480
Qy	482	EGALKLKEISYIHAEAYAAGELKHGPLALIDADMPVIVVAPNPNELLEKLKSNIEEVRARG	541
Db	481	EGALKLKEISYIHAEAYAAGELKHGPLALIDADMPVIVVAPNPNELLEKLKSNIEEVRARG	540
Qy	542	GQLYVFADQDAGFVSSDNMHIIEMPHVEEVIAPIFYTVPLOLLAYHVHVALIKGTDVDQPRN	601
Db	541	GLLYVFADQDAGFTDSEGMKIIQLPHVEEIIAPIFYTVPLOLLSYHVHVALIKGTDVDQPRN	600
Qy	602	LAKSVTVE	609
Db	601	LAKSVTVE	608

Search completed: June 14, 2006, 15:27:41  
Job time : 308.293 secs

**THIS PAGE IS BLANK**



GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 14, 2006, 15:21:37 ; Search time 41.2344 Seconds  
(without alignments)  
1421.048 Million cell updates/sec

Title: US-10-612-779-6  
Perfect score: 3089  
Sequence: 1 MCGIVGAIQRDVAEILLEG.....LIKGTVDQPRNLAKSVTYE 609

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 80:\*  
1: Pirl:\*  
2: Pirl:\*  
3: Pirl:\*  
4: Pirl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3068	99.3	609	1	XNECGM
2	3064	99.2	609	2	H86058
3	3057	99.0	609	2	G91212
4	3035	98.3	609	2	AB0955
5	2668	86.4	609	2	AB0500
6	2256.5	73.0	610	2	D64067
7	2250.5	72.9	610	2	B82316
8	1983	64.2	611	2	F82951
9	1856	60.1	621	2	A84933
10	1688	54.6	635	2	E71272
11	1681	54.4	609	2	A82844
12	1579	51.1	611	2	T45493
13	1569.5	50.8	612	2	H82022
14	1565.5	50.7	612	2	B81246
15	1435	46.5	612	2	AD3595
16	1417.5	45.9	608	2	B95322
17	1413.5	45.8	608	2	C97575
18	1413.5	45.8	608	2	AC2796
19	1404.5	45.5	604	2	S67993
20	1362.5	44.1	606	2	F75536
21	1350.5	43.7	592	2	D70327
22	1333.5	43.2	608	1	S01040
23	1323.5	42.8	606	2	B87263
24	1321	42.8	606	2	S16561
25	1320.5	42.7	598	2	F75212
26	1303.5	42.2	606	2	B72412
27	1287.5	41.7	608	2	C96919
28	1284	41.6	615	2	T35569
29	1282.5	41.5	601	2	D71248

RESULT 1

XNECGM

glutamine-fructose-6-phosphate transaminase (isomerizing) (EC 2.6.1.16) - Escherichia coli  
N;Alternate names: glucosamine-6-phosphate synthase; glucosaminophosphate isomerase (glu  
C;Species: Escherichia coli  
C;Date: 31-Mar-1990 #sequence\_revision 17-Oct-1997 #text\_change 09-Jul-2004  
R;Accession: B65176; A30389; I41219; S17839; Q90513  
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co  
A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997

A;Title: The complete genome sequence of Escherichia coli K-12.  
A;Reference number: A64720; MUID:97426617; PMID:9278503

A;Accession: B65176  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-609 <BLAT>

A;Cross-references: UNIPROT:P17169; UNIPARC:UPI0000153CA3; GB:AE000450; GB:U00096; NID:9  
A;Experimental source: strain K-12, substrain MG1655  
R;Walker, J.E.; Gay, N.J.; Saraste, M.; Eberle, A.N.  
Biochem. J. 224, 799-815, 1984

A;Title: DNA sequence around the Escherichia coli unc operon. Completion of the sequence  
A;Reference number: A30389; MUID:85121806; PMID:6395859

A;Accession: A30389

A;Molecule type: DNA

A;Residues: 1-418,'NV',421-609 <WAL>

A;Cross-references: UNIPARC:UPI000016F5F4; GB:X01631; NID:G43256; PIDN:CAA25785.1; PID:9  
R;McKown, R.L.; Orle, K.A.; Chen, T.; Craig, N.L.  
J. Bacteriol. 170, 352-358, 1988

A;Title: Sequence requirements of Escherichia coli attTn7, a specific site of transposon  
A;Reference number: I41219; MUID:88086894; PMID:2826397

A;Accession: I41219

A;Status: translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 597-609 <RES>  
A;Cross-references: UNIPARC:UPI000016F18D; GB:M1980; NID:G146074; PIDN:AAA23836.1; PID:9  
R;Golinski-Pimpaneau, B.; Badet, B.  
Eur. J. Biochem. 201, 175-182, 1991

A;Title: Possible involvement of Lys603 from Escherichia coli glucosamine-6-phosphate syn  
A;Reference number: S17839; MUID:92007872; PMID:1915361

A;Accession: S17839

A;Molecule type: protein

A;Residues: 49-50,'X',52-53;218;219,220-223,'X',225-231;489-493;504-508,'K','K',601-609-  
A7  
A;Cross-references: UNIPARC:UPI000011EBBE; UNIPARC:UPI000011ECBE; UNIPARC:UPI000011ECBE;  
C;Genetics:

A;Gene: glmS

A;Map position: 4 min

C;Function:

A;Description: catalyzes the formation of D-glucosamine-6-phosphate from the amido group  
A;Pathway: glucosamine biosynthesis

A;Note: Glucosamine-6-phosphate is used in the biosynthesis of amino sugars of asparagin

C:Superfamily: glutamine-fructose-6-phosphate aminotransferase (isomerizing)  
C:Keywords: aminotransferase; isomerase  
F:2-609/Product: Glutamine-fructose-6-phosphate transaminase (isomerizing) #status predicted  
F:2/Active site: Cys #status predicted

Query Match 99.3%; Score 3068; DB 1; Length 609;  
Best Local Similarity 99.5%; Pred. No. 6e-191;  
Matches 606; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MCGIVGATAQRDVAEILLEGRLRLEYRGYDVSAGLAVVDTEGHMTLRLRLGKQVQMLAQAAE 60  
DB 1 MCGIVGATAQRDVAEILLEGRLRLEYRGYDVSAGLAVVDTEGHMTLRLRLGKQVQMLAQAAE 60

QY 61 EPHLGGTGIAHTRWATHGEPSEVNAPHVSHIIVVHNGIITENHEPLREELKARGYTFV 120  
DB 61 EPHLGGTGIAHTRWATHGEPSEVNAPHVSHIIVVHNGIITENHEPLREELKARGYTFV 120

QY 121 SETDTEVIAHLVNWELKOGGTLREAVLRAIPOLRGAYGTVIMDSRHPDTLLAARSGSPV 180  
DB 121 SETDTEVIAHLVNWELKOGGTLREAVLRAIPOLRGAYGTVIMDSRHPDTLLAARSGSPV 180

QY 181 IGLGNGENFIASDQALLPVTRRFIFLEEGDIAETRRSVNIFDKTGAEVKQDIESNLQ 240  
DB 181 IGLGNGENFIASDQALLPVTRRFIFLEEGDIAETRRSVNIFDKTGAEVKQDIESNLQ 240

QY 241 YDAGDKGIYCHYMOKEIYEQPNAINKNTLTGRISHGOVDLSELGPNADELLSKVEHIQIIA 300  
DB 241 YDAGDKGIYCHYMOKEIYEQPNAINKNTLTGRISHGOVDLSELGPNADELLSKVEHIQIIA 300

QY 301 CGTSYNSGMVSRYPFESLAGIPCDEVIASEPRYKSAVRNLSLMTTLSQSGTADTLA 360  
DB 301 CGTSYNSGMVSRYPFESLAGIPCDEVIASEPRYKSAVRNLSLMTTLSQSGTADTLA 360

QY 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQTLTVLLMLVAKL 420  
DB 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQTLTVLLMLVAKL 420

QY 421 SRLKGLDASIEHDIHVHGLQALPSRIEQMSODKRIEALAEAFSDKHAFFLSRGQOYPIA 480  
DB 421 SRLKGLDASIEHDIHVHGLQALPSRIEQMSODKRIEALAEAFSDKHAFFLSRGQOYPIA 480

QY 481 LEGALKKEISYIHAEVAAAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEVRAR 540  
DB 481 LEGALKKEISYIHAEVAAAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEVRAR 540

QY 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPIFTYVPLQLLAYHVALIKGTDVDDQPR 600  
DB 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPIFTYVPLQLLAYHVALIKGTDVDDQPR 600

QY 601 NLAKSVTVE 609  
DB 601 NLAKSVTVE 609

RESULT 2  
H86058  
hypothetical protein glms [imported] - Escherichia coli (strain O157:H7, substrain EDL93)  
C:Species: Escherichia coli  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 09-Jul-2004  
C:Accession: H86058  
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glaener, J.D.; Rose, D.J.; Mayhew  
iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouais, K.; Apodaca,  
Nature 409, 529-533, 2001  
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A:Reference number: A85480; MUID:21074935; PMID:11206551  
A:Accession: H86058  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-609 <STO>  
A:Cross-references: UNIPROT:Q8XEG2; UNIPARC:UPI0000165991; GB:AE005174; NID:g12518583; F  
C:Experimental source: strain O157:H7, substrain EDL933  
C:Genetics:  
A:Gene: glms

C:Superfamily: glutamine-fructose-6-phosphate aminotransferase (isomerizing)  
C:Keywords: aminotransferase; isomerase  
F:2-609/Product: Glutamine-fructose-6-phosphate transaminase (isomerizing) #status predicted  
F:2/Active site: Cys #status predicted

Query Match 99.2%; Score 3064; DB 2; Length 609;  
Best Local Similarity 99.3%; Pred. No. 1.1e-190;  
Matches 605; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MCGIVGATAQRDVAEILLEGRLRLEYRGYDVSAGLAVVDTEGHMTLRLRLGKQVQMLAQAAE 60  
DB 1 MCGIVGATAQRDVAEILLEGRLRLEYRGYDVSAGLAVVDTEGHMTLRLRLGKQVQMLAQAAE 60

QY 61 EPHLGGTGIAHTRWATHGEPSEVNAPHVSEHIIVVHNGIITENHEPLREELKARGYTFV 120  
DB 61 EPHLGGTGIAHTRWATHGEPSEVNAPHVSEHIIVVHNGIITENHEPLREELKARGYTFV 120

QY 121 SETDTEVIAHLVNWELKOGGTLREAVLRAIPOLRGAYGTVIMDSRHPDTLLAARSGSPV 180  
DB 121 SETDTEVIAHLVNWELKOGGTLREAVLRAIPOLRGAYGTVIMDSRHPDTLLAARSGSPV 180

QY 181 IGLGNGENFIASDQALLPVTRRFIFLEEGDIAETRRSVNIFDKTGAEVKQDIESNLQ 240  
DB 181 IGLGNGENFIASDQALLPVTRRFIFLEEGDIAETRRSVNIFDKTGAEVKQDIESNLQ 240

QY 241 YDAGDKGIYCHYMOKEIYEQPNAINKNTLTGRISHGOVDLSELGPNADELLSKVEHIQIIA 300  
DB 241 YDAGDKGIYCHYMOKEIYEQPNAINKNTLTGRISHGOVDLSELGPNADELLSKVEHIQIIA 300

QY 301 CGTSYNSGMVSRYPFESLAGIPCDEVIASEPRYKSAVRNLSLMTTLSQSGTADTLA 360  
DB 301 CGTSYNSGMVSRYPFESLAGIPCDEVIASEPRYKSAVRNLSLMTTLSQSGTADTLA 360

QY 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQTLTVLLMLVAKL 420  
DB 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQTLTVLLMLVAKL 420

QY 421 SRLKGLDASIEHDIHVHGLQALPSRIEQMSODKRIEALAEAFSDKHAFFLSRGQOYPIA 480  
DB 421 SRLKGLDASIEHDIHVHGLQALPSRIEQMSODKRIEALAEAFSDKHAFFLSRGQOYPIA 480

QY 481 LEGALKKEISYIHAEVAAAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEVRAR 540  
DB 481 LEGALKKEISYIHAEVAAAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEVRAR 540

QY 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPIFTYVPLQLLAYHVALIKGTDVDDQPR 600  
DB 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPIFTYVPLQLLAYHVALIKGTDVDDQPR 600

QY 601 NLAKSVTVE 609  
DB 601 NLAKSVTVE 609

RESULT 3  
G91212  
hypothetical protein ECs4671 [imported] - Escherichia coli (strain O157:H7, substrain RIN  
C:Species: Escherichia coli  
C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 09-Jul-2004  
C:Accession: G91212  
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;  
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genom  
A:Reference number: A99629; MUID:21156231; PMID:11258796  
A:Accession: G91212  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-609 <HAY>  
A:Cross-references: UNIPROT:Q8XEG2; UNIPARC:UPI000016552F; GB:BA000007; PIDN:BAE38094.1;  
A:Experimental source: strain O157:H7, substrain RIND 0509952  
C:Genetics:  
A:Gene: ECs4671  
C:Superfamily: glutamine-fructose-6-phosphate aminotransferase (isomerizing)  
Query Match 99.0%; Score 3057; DB 2; Length 609;

	Best Local Similarity	99.2%; Pred. No. 3.le-190;	Matches	604; Conservative	2; Mismatches	3; Indels	0; Gaps	0;
Qy	1	MCGIVGAIAORDVAEILLEGRLRLEYRGYDSAGLAVVDTEGHMTLRRLRKGVQMLAQAAE	60					
Db	1	MCGIVGAIAORDVAEILLEGRLRLEYRGYDSAGLAVVDAEGHMTLRRLRKGVQMLAQAAE	60					
Qy	61	EHPHLGGTGTAHTRWATHGPSEPVNAHPHVSEHIIVVHNGIIENHEPULREELKARGYTFV	120					
Db	61	EHPHLGGTGTAHTRWATHGPSEPVNAHPHVSEHIIVVHNGIIEHNPEULREELKARGYTFV	120					
Qy	121	SETDTEVIAHLVNVELKOGGTFLREAVLRAIPQLRGAYGTVMDSRHDPDTLLAARSGSLV	180					
Db	121	SETDTEVIAHLVNVELKOGGTFLREAVLRAIPQLRGAYGTVMDSRHDPDTLLAARSGSLV	180					
Qy	181	IGLGMGNFASDLQALLPVTRRFIFLEEGDIAETIRRSVNIFDKTGAEVKRQDIESNLQ	240					
Db	181	IGLGMGNFASDLQALLPVTRRFIFLEEGDIAETIRRSVNIFDKTGAEVKRQDIESNLQ	240					
Qy	241	YDAGDKGIYCHYMOKEIYEOPNAIKNTLTGRISHGOVDLSBELGNABELLSKVBIHQILA	300					
Db	241	YDAGDKGIYRHYYMOKEIYEOPNAIKNTLTGRISHGOVDLSBELGNABELLSKVBIHQILA	300					
Qy	301	CGTSYNSGMWSRYWFESIAGIPCDEVTASEPRYKSAVRNRSLMITLSOSGETADTLAGL	360					
Db	301	CGTSYNSGMWSRYWFESIAGIPCDEVTASEPRYKSAVRNRSLMITLSOSGETADTLAGL	360					
Qy	361	RLSKELGYLGS LAICNVPGSSLVRES DLALM TNAGTEIGVASTKAFTTQLTVLLMLVAKL	420					
Db	361	RLSKELGYLGS LAICNVPGSSLVRES DLALM TNAGTEIGVASTKAFTTQLTVLLMLVAKL	420					
Qy	421	SRLKGDLAS IEHDIVHG LQALPSRIEOWLSODKRIEALAEAFDSKHHALFISRGDQVP IA	480					
Db	421	SRLKGDLAS IEHDIVHG LQALPSRIEQMLSODKRIEALAEAFDSKHHALFISRGDQVP IA	480					
Qy	481	LEGALKLKEI SYTHAEAYAAGELKHGPLALIDAMPVTIVAPNNELLEKLKSNIIEEVAR	540					
Db	481	LEGALKLKEI SYTHAEAYAAGELKHGPLALIDAMPVTIVAPNNELLEKLKSNIIEEVAR	540					
Qy	541	GGOLYVFADQDAGCVSSDNMH IIEMPHVEEVIAP IFYTPVQLLAYHVHVALIKGTDVDQPR	600					
Db	541	GGOLYVFAEQDAGCVSSDNMH IIEMPNVEEVIAP IFYTPVQLLAYHVHVALIKGTDVDQPR	600					
Qy	601	NLAKSVTVE	609					
Db	601	NLAKSVTVE	609					

## RESULT 4

AB0955  
Glutamine-fructose-6-phosphate transaminase (isomerizing) (EC 2.6.1.16) - *Salmonella enteritidis*  
N;Alternate names: glucosamine-fructose-6-P aminotransferase [misnomer]  
C;Species: *Salmonella enterica* subsp. *enterica* serovar Typhi  
A;Note: this species has also been called *Salmonella typhi*  
C;Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 28-Jul-2003  
C;Accession: AB0955  
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, K.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, P.; S.; Moule, S.; O'Gaora, P.  
Nature 413, 848-852, 2001  
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; A;Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Typhi  
A;Reference number: AB0502; MUID:21534947; PMID:11677608  
A;Accession: AB0955  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-609 <PAR>  
A;cross-references: UNIPARC:UPT000005A683; GB:AL513382; PIDN:CAD03134.1; PID:g16504770;  
C;Genetics:  
A;Gene: STY3917  
C;Superfamily: Glutamine-fructose-6-phosphate aminotransferase (isomerizing)  
C;Keywords: aminotransferase; isomerase

Query Match	98.3%;	Score 3035;	DB 2;	Length 609;
Best Local Similarity	98.0%;	Pred. No. 8.2e-189;		
Matches 597;	Conservative 5;	Mismatches 7;	Indels 0;	Gaps 0;
Qy	1	MCGIVGAIAORDVAEILLEGRLREYRGYDSAGLAVVDTTECHMTRLRRLRGKVOMLAQAAE	60	
Db	1	MCGIVGAIAORDVAEILLEGRLREYRGYDSAGLAVVDAEGHMTLRRLRGKVOMLAQAAE	60	
Qy	61	EHPHGGTGTAHTRWATHGEPSEVNAHPHVSEHIVVHNGIIENHEPLREELKARGYTFV	120	
Db	61	EHPHGGTGTAHTRWATHGEPSEANAPHVSEHIVVHNGIIENHEPLREALKARGYTFV	120	
Qy	121	SETDTEVIAHLVNWELKQGGTLEAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGSPLV	180	
Db	121	SETDTEVIAHLVNWELKQGGTLDAILRAIPQLRGAYGTVMDSRHPDITLLAARSGSPLV	180	
Qy	181	IGLWGNFIASDQALLPVTRRPIFLEEGDIAEITRRSVNI PDKTGAEVKRDIESNLQ	240	
Db	181	IGLWGNFIASDQALLPVTRRPIFLEEGDIAEITRRSVNI PDKTGAEVKRDIESNLQ	240	
Qy	241	YDAGDKGIGYCHYMKETIEQPNAIKNTLGRISHGQVDLSGLGNADLLSKVEHIOILA	300	
Db	241	YDAGDKGIGYRHVMKETIEQPNAIKNTLGRISHGQVDLSGLGNADLLSKVEHIOILA	300	
Qy	301	CGTSYNSGWSRYWFESLAGIPCDVEIASEFRYKSAVRNSLMI TLSQSGETADTTLAGL	360	
Db	301	CGTSYNSGWSRYWFESLAGIPCDVEIASEFRYKSAVRNSLMI TLSQSGETADTTLAGL	360	
Qy	361	RLSKELGYLGS LAICNVPGSSLRRESDLALMTNAGTEIGVASTKAFPTQITVLLMLVAKL	420	
Db	361	RLSKELGYLGS LAICNVPGSSLRRESDLALMTNAGTEIGVASTKAFPTQITVLLMLVAKL	420	
Qy	421	SRLKGLDASIEHDIVHGLQALPSRIEQMLSQDKRIEALAE DFD SKH ALFLSRGDQYPIA	480	
Db	421	SRLKGLDASIEHDIVHGLQALPSRIEQMLSQDKRIEALAE DFD SKH ALFLGRGDQYPIA	480	
Qy	481	LEGALKUKETISYIHA EYAAAGELKHGPLALIDAMPVIVVAPNNELLEKLKSNIEEVRAR	540	
Db	481	LEGALKUKETISYIHA EYAAAGELKHGPLALIDAMPVIVVAPNNELLEKLKSNIEEVRAR	540	
Qy	541	GGQLYVFADDDAGVSSDNNHHIIT E MPHVEEVIAP I FVTVPQLL LAYHVALIKGTDVDPQR	600	
Db	541	GGQLYVFSDQDAGVSSDNNHHIIT E MPHVEEVIAP I FVTVPQLL LAYHVALIKGTDVDPQR	600	
Qy	601	NLAKSVTVE	609	
Db	601	NLAKSVTVE	609	

## RESULT 5

AB05000  
glutamine-fructose-6-phosphate transaminase (isomerizing) (EC 2.6.1.16) [imported] - Ver  
C.Species: *Yersinia pestis*  
C.Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Jul-2004  
C.Accession: AB05000  
R.Paper: Hall, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.  
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;  
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,  
Nature 413, 523-527, 2001  
A.Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.  
A.Reference number: AB00001; PMID:21470413; PMID:11586360

Query Match 86.4%; Score 2668; DB 2; Length 609;  
Best Local Similarity 84.7%; Pred. No. 4.9e-165;

Matches 516; Conservative 44; Mismatches 49; Indels 0; Gaps 0;

QY 1 MCGIVGAIQORDVAEILLEGRLREYRGYDSAGLAVVDTEGHMTLRRLRGKVMQAQAAE 60  
 Db 1 MCGIVGAIQORDVAEILLEGRLREYRGYDSAGLAVVDTEGHMTLRRLRGKVMQAQAAE 60  
 QY 61 EHPHGGTGIAHTRWATHGSPSEVNAHPHSEHIVVYVHNGIIENHPELREELKARGYTFV 120  
 Db 61 KQDLHGCTGIAHTRWATHGSPSEANAHPHSDVSVVHNGIIENHPELRELLISRGVRF 120  
 QY 121 SETDTEVIAHLVNWELKQGTLEAVLRATPOLRGAYGTVMDSRHPDPTLLAARSGSPLV 180  
 Db 121 SETDTEVIAHLVNWEEQGGSLLEVKRVIPOLRGAYGTVMDSRPSRLIAARSGSPLV 180  
 QY 181 IGLGMGNFTASDQALLPVTFRPIFLEEGDIAETRRSVNIPDKTGAEVKRODIESNLQ 240  
 Db 181 IGGVGENFTASDQALLPVTFRPIFLEEGDVVEVTRRSISIFDKQNAIERPEIESQVQ 240  
 QY 241 YDAGDKGIYCHYMQEYEQFNAIKNTLTGRISHGQVDLSSELGNADLLSKVEHIQILA 300  
 Db 241 YDAGDKGIYRHYMQEYEQFMAIKNTLTGRLSHGMDLSLGPADALLAEVQHIQIIA 300  
 QY 301 CGTSYNSGMVSRWTFESLAGIPDVEITASFYRKSAVRNSLMITLSQSGETADTLA 360  
 Db 301 CGTSYNSGMVSRWTFESLAGIPDVEITASFYRKSAVRNSLLITLSQSGETADTLAAL 360  
 QY 361 RLSKELGYLGSALAINVPGSSLVRESDLMTWAGTEIGVASTKFTTQTLTLLMLVAKL 420  
 Db 361 RLSKELGYLGSALAINVAGSSLVRESDLMTWAGTEIGVASTKFTTQTLTLLMLVGR 420  
 QY 421 SRLKGLDASTEHIDVHGLQALPSRIEQLSQDKRIEALAEFSDKHGHALFSLRGDQVPIA 480  
 Db 421 GKLGKADASLEHIDVHALQALPARIEQLSLDKTIEALAEFSDKHGHALFLRGDQVPIA 480  
 QY 481 LEGALKUKETSYTHAEYAAAGELKHGPLALIDAMPVIVVAPNNELLEKKUSNIEVRAR 540  
 Db 481 MEGALKUKETSYTHAEYAAAGELKHGPLALIDAMPVIVVAPNNELLEKKUSNIEVRAR 540  
 QY 541 GGLLYVPADODAGVSSDNNHIIEMPHVEVPIAFYVTPQLQLAHVALIKGTVDVQPR 600  
 Db 541 GGLLYVPADODAGTDSGKWKIQLPHVEVPIAFYVTPQLQLAHVALIKGTVDVQPR 600  
 QY 601 NLAKSVTVE 609  
 Db 601 NLAKSVTVE 609

RESULT 6  
 D64067  
 glutamine-fructose-6-phosphate transaminase (isomerizing) (EC 2.6.1.16) - Haemophilus influenzae  
 C/Species: Haemophilus influenzae  
 C/Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 09-Jul-2004  
 C/Accession: D64067  
 R/Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, P.A.;  
 Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.;  
 D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.  
 Science 269, 496-512, 1995  
 A/Authors: Gnehm, C.I.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A.;  
 A/Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.  
 A/Reference number: A64000; MUID:95350630; PMID:7542800  
 A/Accession: D64067  
 A/Status: nucleic acid sequence not shown; translation not shown  
 A/Molecule type: DNA  
 A/Residues: 1-610 <TIGR>  
 A/Cross-references: UNIPROT:P44708; UNIPARC:UPI00000512BF; GB:U32726; GB:I42023; NID:G1S  
 A/Note: named as homolog to a protein from Escherichia coli  
 C/Superfamily: glutamine-fructose-6-phosphate aminotransferase (isomerizing)  
 C/Keywords: aminotransferase; isomerase  
 F/2-610/Product: glutamine-fructose-6-phosphate transaminase (isomerizing) #status predicted  
 F/2-Active site: Cys #status predicted

Query Match 73.0%; Score 2256.5; DB 2; Length 610;  
 Best Local Similarity 72.0%; Pred. No. 2.2e-138;

Matches 439; Conservative 73; Mismatches 97; Indels 1; Gaps 1;

```

QY 1 MCGIVGAIAQRDVAEIILEGRLREYRGYDSAGLAVVDTEGHMTRRLRLGKVMQLAAAE 60
Db 1 MCGIVGVAQRDVAEIIINGLHRLREYRGYDSAGVAVINKQELQRIICLQKVKALDEAVS 60
QY 61 EPHLHGSGTGAHTRWATHGBEPSEVNAHPHVSHEHVVVHNGIENHEPELREELKARGYTFV 120
Db 61 EKPLIGGTGAHTRWATHGBEPSETNAHPHSSGTFAVHNGIENHEELRELLKSRGYVFL 120
QY 121 SETDTEVIAHLVNWELKQGTLEAVLRAIPOLRGAYCTVIMDSRHPDPTLLAARSGSPLV 180
Db 121 SQDTEVIAHLVNWEMKTTQSLDVAVKAKQLTGAYGVVMSRHPDPTLLAARSGSPLV 180
QY 181 IGLGMGNFTASQALLLPVTTRFIFLEEGDIAEITRRSVNIFDKTGAEVKRDIESNLQ 240
Db 181 IGLGIGENFLASQALLLSVTRFIFLEEGDIAEITRTVDIYDTHGNKAKREIHESNLE 240
QY 241 YDAGDKGIYCHYMQKEIYEQNAIKNTLTGRISHGQVDLSLGPNADELLEKVEHIQILA 300
Db 241 NDAAEKGRFHFQKEIYEQPTALINTMEGRINHENVIVDSIGNKAGKILEKVEHIQIVA 300
QY 301 CGTSYNSGMVSRVWFESLAGIPCDVETASBPYRKSAVRNLSMTITLSQSGETADTLACL 360
Db 301 CGTSYNSGMVSRVWFESLAGVSCDVEIASBPYRKSVTRPNSLITITLSQSGETADTLAAL 360
QY 361 RLSKELGYLGLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFTTQLTVLLMLVAKL 420
Db 361 FLAKKGYMAALTICNVAGSSLVRESDLAPMTRAGVGVGASTKAFTTQLAALLMLVTAL 420
QY 421 SRLKG-LDASIEHDIVHGLQALPSRIQMSODKRIEALAEDEFSDKHHALFLSRGDQYPI 479
Db 421 GKVKGHISVEKEREIIKAMQSLPAEIEKALAFOTEIEALAEDEFAEKHHALFLGRGAFYPI 480
QY 480 ALGALKLKEISYTHAEYAAAGELKHGPLALIDADMPVIVVAPNNLELLEKKSNEEVRA 539
Db 481 AVFASLKLKEISYTHAEYAAAGELKHGPLALIDADMPVIVVAPNNLELLEKKSNEEVRA 540
QY 540 RGGQLYVFADQDAGFVSSDNHIIEMPHVBEVIAPIFYTVPLQLLAYHVALIKGTDVDP 599
Db 541 RGGQLYVFADKEAGFTSEGMKIITMPKVDIVAPIFYTIPMQLLIYYVALIKGTDVDP 600
QY 600 RNLAKSVTVE 609
Db 601 RNLAKSVTVE 610

```

RESULT 7  
E82316  
glutamine-fructose-6-phosphate transaminase (isomerizing) (EC 2.6.1.16) [similarity] - V3  
C:Species: Vibrio cholerae  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004  
C:Accession: E82316  
R:Heidelberg, J.F.; Eisen, J.D.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;  
chardson, D.; McInerney, J.A.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.  
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A:Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.  
A:Reference number: A82035; MUID:20406833; PMID:10952301  
A:Accession: E82316  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-610 <HEI>  
A:Cross-references: UNIPROT:Q9KUM8; UNIPARC:UPI0000164B53; GB:AE004135; GB:AE003852; NID:  
A:Experimental source: serogroup O1; strain N16961; biotype El Tor  
C:Genetics:  
A:Gene: VC0487  
A:Map position: 1  
C:Superfamily: glutamine-fructose-6-phosphate aminotransferase (isomerizing)  
C:Keywords: aminotransferase; isomerase  
F:2-610/Product: glutamine-fructose-6-phosphate transaminase (isomerizing) #status predicted  
F:2/Active site: Cys #status predicted

Query Match	73.0%;	Score 2256.5;	DB 2;	Length 610;
Best Local Similarity	72.0%;	Pred. No. 2.2e-138;		

Query Match 72.9%; Score 2250.5; DB 2; Length 610;  
Best Local Similarity 71.3%; Pred. No. 5.4e-138; Mismatches 102; Indels 1; Gaps 1;  
Matches 435; Conservative 77;

1 MCGIVGAIQORDVAEILLEGRLRLEYRGYDSAGLAVVDTTEGHMTRRLRLKGVQMLAQAAE 60  
1 MCGIVGAVQORDVAEILVQGLRLEYRGYDSAGVAVVDSKQLTRRLKGVQELADAVE 60

61 EHPHGGTGTIAHTRWATHGEPSEVNAHPHV-SHIVVHNGIIEHHEPLEELKARGYTFV 120  
61 AAQVAGGTGTIAHTRWATHGEPSEVNAHPHISGDIIVVHNGIIEHHEMLRMLQDQGVFT 120

121 SETDTEVIAHLVNMELKQGGTLEAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGSPLV 180  
121 SQTDEVIHLVNMELKQGGTLEAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGSPLV 180

181 IGLGMENFIASDQALLPVTRRFIFLEEGDIAEITRRSVNIFDKTGAEVKQDIESNLQ 240  
181 IGFIGENFLASDQALLPVTRRFIFLEEGDIAEITRRSVNIFDKTGAEVKQDIESNLQ 240

241 YDAGDKGIYCHYNQKEIYEOPNAIKNTLTGRISHGQVDLSELGPNADLLSKVEHTQILA 300  
241 HDAADKGYRHYMKQKEIFEQPKALINTMEGRITHDCCVVVESIGVHAAEILAKVEHVQIVA 300

301 CGTSYNSGMVSRVWFESLAGIPCDVEIASFPRYKSAVRNLSMITLSQSGETADTLAAGL 360  
301 CGTSYNAGMTARVWFESLAGVCDVEIASFPRYKSAVRNLSMITLSQSGETADTLAAL 360

361 RLKSELGYLGSALAI CNVPGSSLVRESDLALMTNAGTIGVASTKAFPTTQTLVLLMLVAKL 420  
361 RLAKENGYMAAMTICNVAGSSLVRESDFAPWTRAGTIGVASTKAFPTTQTLVLLMLVAKL 420

421 SR-LKGLDASIEDHIVHGLQALPSRIEOMLSQDKRIEALAEPSDKHAFILSRGQOYPI 479  
421 GKQOQIRIGREAEIVHALQPKQIETALSFEKQIETLAEDFADKHHTFLGRGEYPI 480

480 ALLEGALKKEISYTHAEAYAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEBVR 539  
481 AVBASLKLKEISYTHAEAYAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEBVR 540

540 RGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPFYTPVPLQLLAYHVALIKGTVDQDP 599  
541 RGGLYVFADQDAGFVSSDNMHIEMPHVEEVIAPFYTPVPLQLLAYHVALIKGTVDQDP 600

600 RNLAKESTVVE 609  
601 RNLAKESTVVE 610

RESULT 8  
F82951  
glutamine-fructose-6-phosphate transaminase (isomerizing) (EC 2.6.1.16) [similarity] - P  
C:Species: Pseudomonas aeruginosa  
C>Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004  
C:Accession: F82951  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B.  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho  
A:Reference number: A82950; MUID:20437337; PMID:10984043  
A:Accession: F82951  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-611 <STO>  
A:Cross-references: UNIPROT:Q9HT25; UNIPARC:UPI0000165E0; GB:AE004091; NID  
A:Experimental source: strain PAO1  
C:Genetics:  
A:Gene: glms; PA5549  
C:Superfamily: Glutamine-fructose-6-phosphate aminotransferase (isomerizing)  
C:Keywords: aminotransferase; isomerase  
F;2-611/Product: glutamine-fructose-6-phosphate transaminase (isomerizing) #status predi  
F;2/Active site: Cys #status predicted

Query Match 64.2%; Score 1983; DB 2; Length 611;  
Best Local Similarity 62.4%; Pred. No. 1.2e-120; Mismatches 127; Indels 2; Gaps 2;  
Matches 381; Conservative 101;

1 MCGIVGAIQORDVAEILLEGRLRLEYRGYDSAGLAVVDTTEGHMTRRLRLKGVQMLAQAAE 60  
1 MCGIVGAIQORDVAEILLEGRLRLEYRGYDSAGVAVVDSKQLTRRLKGVQELADAVE 60

61 EHPHGGTGTIAHTRWATHGEPSEVNAHPHV-SHIVVHNGIIEHHEPLEELKARGYTFV 119  
61 GTPLGLGLGTIAHTRWATHGEPSEVNAHPHISGDIIVVHNGIIEHHEPLEERLKLGLYVF 120

120 VSETDTEVIAHLVNMELKQGGTLEAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGSPL 179  
121 TSQDTEVIAHLVNMELKQGGTLEAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGSPL 180

180 VIGLGMENFIASDQALLPVTRRFIFLEEGDIAEITRRSVNIFDKTGAEVKQDIESNL 239  
181 VIGLGMENFIASDQALLPVTRRFIFLEEGDIAEITRRSVNIFDKTGAEVKQDIESNL 240

240 QYDAGDKGIYCHYNQKEIYEOPNAIKNTLTGRISHGQVDLSELGPNADLLSKVEHTQILA 299  
241 GAEADKGYRHYMKQKEIFEQPKALINTMEGRITHDCCVVVESIGVHAAEILAKVEHVQIVA 300

300 ACCTSNSGMVSRVWFESLAGIPCDVEIASFPRYKSAVRNLSMITLSQSGETADTLAAGL 359  
301 ACCTSNSGMVSRVWFESLAGIPCDVEIASFPRYKSAVRNLSMITLSQSGETADTLAAL 360

360 LRLKSELGYLGSALAI CNVPGSSLVRESDLALMTNAGTIGVASTKAFPTTQTLVLLMLVAK 419  
361 LRNAKELGFLSSVAICNVATSSLVRESDLTLLTAQAGPEIGVASTKAFPTTQTLVLLMLVAK 420

420 LSRL-LKGLDASIEDHIVHGLQALPSRIEOMLSQDKRIEALAEPSDKHAFILSRGQOY 478  
421 IGQVQKRLGADGVAELVDELRLPTLGEALAMNRTVEKVSSELFAEKHHTFLGRGAQFP 480

479 IALEGALKKEISYTHAEAYAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEBVR 538  
481 VALEGALKKEISYTHAEAYAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEBVR 540

539 ARGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPFYTPVPLQLLAYHVALIKGTVDQDP 598  
541 ARGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPFYTPVPLQLLAYHVALIKGTVDQDP 600

599 PRNLAKESTVVE 609  
601 PRNLAKESTVVE 611

RESULT 9  
A84933  
glutamine-fructose-6-phosphate transaminase (isomerizing) (EC 2.6.1.16) [imported] - Buch  
N;Alternate names: D-fructose-6-phosphate amidotransferase  
C:Species: Buchnera sp.  
C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 28-Jul-2003  
C:Accession: A84933  
R:Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.  
Nature 407, 81-86, 2000  
A:Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. A  
A:Reference number: A84930; MUID:20445173; PMID:10993077  
A:Accession: A84933  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-621 <STO>  
A:Cross-references: UNIPARC:UPI000005E424; GB:AP000398; GSPDB:GN00144  
A:Experimental source: strain APS  
C:Genetics:  
A:Gene: glms; BU026  
C:Superfamily: Glutamine-fructose-6-phosphate aminotransferase (isomerizing)  
C:Keywords: aminotransferase; isomerase  
Query Match 60.1%; Score 1856; DB 2; Length 621;

		Best Local Similarity 58.8%; Pred. No. 28-112; Matches 359; Conservative 116; Mismatches 132; Indels 4; Gaps 3;			
QY	1	MCGIVGAIQAORDVAEILLEGRLREYRGYDSAGLAVVDTEGHMTRLRLRLGKQVQLAQAAE	60		
Db	13	MCGIVAAVTQRIANFLIDGKKLEVRGYDSSGLAVINDKNKNIVRICRGKVNELIKKTN	72		
QY	61	EHPHGGGTGIAHTRWATHGEPSEVNAHPVSHIIVVHNGIENHEPRLREELKARGYTFV	120		
Db	73	KKKILGSGIAGVATHWATHGSKENHTPHISSNIIVVHNGIENNSTLGRFLKKQGYIFS	132		
QY	121	SETDTEVIAHLVNWEL-KGGGTLRAVLRAIPQLRGAYGTVIMDSRHPDPTLLAARSGSPL	179		
Db	133	SDTIDEVIAHLHWELQKKQKSLIKVIQNSIKKLGNSVMVVIDQNPKLTAARSGSPL	192		
QY	180	VIGLGMGENFIASDQALLPVTRRFIFLEEGDIAETRRSVNIFDKTGAEVKRODIESNL	239		
Db	193	IIGLGTENFIASDQIALAHVTKRFIYLEEGDIAIVARKEINI FNKNSIIQREEVVSN	252		
QY	240	QYDAGDKGIYCHYMOKEIYEQNAIKNTLTGRIHSH-GQVDLSSELGNADDELLSKVEHIQ	298		
Db	253	EYKSAKGGKRYRMEKEIHEQPKSIRNTLKNRLTNSKNVHFSELGSKENNIFYNTEHIQ	312		
QY	299	LACGTSYNSGMSYRWFEESLAGIPCDVEITASEFRYRKSAAVRENSLMITLSQSGETADTLA	358		
Db	313	VACGTSYNAWVSRVFEELANIPCDVEITASEFSRKLVRKKSLLITLSQSGETADTILS	372		
QY	359	GURLSKELGSLGSLAI CNVPGSSLVRESDLAMTNAGTEIGVASTKAFPTQTLTVLLMLVA	418		
Db	373	ALRYSKGLGYNLITCNMKSSSLVRESDFYILTRAGLEIGVASTKSPFTQTLTVLLMLVA	432		
QY	419	KLSRLKGLDASTEHIDVHGLQALPSRIEQMLSDQKRIEALABDFDKHALPLSRGDQVP	478		
Db	433	KIINGCKENNTSKRIVQTLSPVRIBELKKQLIQDMANTLANKNMFLGRGNQVP	492		
QY	479	ILEGALKLKEISYTHAEAYAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEVR	538		
Db	493	IAMEGALKLKEISYTHAEAYSGELKHGPLALIDKNPIMTAPENSLERKKNKKEIC	552		
QY	539	ARGGOLYFADQDAGFVSSDNHIIEMPHVEVIAPIFYTVPLQLLAYHVALIKGTDVQD	598		
Db	553	SRGGIVVYFSNQE--FDYENINTIKLPYVEELIAPIFYTIPQLFAYYVALKKGRDIDQ	610		
QY	599	PRNLAKSVTVE 609			
Db	611	PHRLAKSVTVE 621			

RESULT 10  
E1272  
Glutamine-fructose-6-phosphate transaminase (isomerizing) (EC 2.6.1.16) glms - syphilis  
C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)  
C:Date: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 09-Jul-2004  
C:Accession: E71272  
R:Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin  
they, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McD  
Science 281, 375-388, 1998  
A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.  
A:Reference number: A71250; MUID:98332770; PMID:9665876  
A:Accession: E71272  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-635 <COL>  
A:Cross-references: UNIPROT:O83833; UNIPARC:UPI0000164A8C; GB:AE001256; GB:AE000520; NID  
A:Experimental source: strain Nichols  
A:Genetics:  
C:Gene: TP0861  
C:Superfamily: glutamine-fructose-6-phosphate aminotransferase (isomerizing)  
C:Keywords: aminotransferase; isomerase  
F:2-635/Product: Glutamine-fructose-6-phosphate transaminase (isomerizing) #status predi  
F:2/Active site: Cys #status predicted

		Query Match 54.6%; Score 1688; DB 2; Length 635; Best Local Similarity 53.7%; Pred. No. 1.6e-101; Matches 341; Conservative 98; Mismatches 170; Indels 26; Gaps 5;			
QY	1	MCGIVGAIQAORDVAEILLEGRLREYRGYDSAGLAVVDTEGHMTRLRLRLGKQVQLAQAAE	60		
Db	1	MCGIVGVAGRDVSGLLLEGRLREYRGYDAGIAGVGSDCALRLRCRCEGRVQSLCALLG	60		
QY	61	EHPHGGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIENHEPRLREELKARGYTFV	120		
Db	61	QSPCLGTGIAHTRWATHGKPCANAHPHCSVAIVHNGIENHRSRLREMLVTRGYFFH	120		
QY	121	SETDTEVIAHLVNWELKOGGTLREAVLRAIPQLRGAYGTVIMDSRHPDPTLLAARSGSPLV	180		
Db	121	SQTDEVIHLHWELRYTAHLHLAVKVLTVQVRGTGLLNCDAASPGRLIAARSGSPLA	180		
QY	181	IGLGMGENFIASDQALLPVTRRFIFLEEGDIAETRRSVNIFDKTGAEVKRODIESNQ	240		
Db	181	VOLGCGENFVTSDDLALAHVTQRFIYLEEGDIADVHRDSVCVHDAQGNVAVRPVTVQMQ	240		
QY	241	YDAGDKGIYCHYMOKEIYEQNAIKNTLTGRISHGQVDLSSEL-----GPN-----D	287		
Db	241	LCTQDKGTHRHMHQEIWQPHAIRHTLNAVMSFSSSRAQVTRFGEDRVLDGTSCKTPE	300		
QY	288	ELLSKVEHIQLACGTSYNSGMSYRWFEESLAGIPCDVEITASEFRYRKSAAVRENSLMITL	347		
Db	301	RLFRITRVRITACGTSYHAGLVARYWFEAFAGVCGQVEIASEYRYSRVHAREIVLTI	360		
QY	348	SQSGETADTLAAGRLSKELGYLSLAI CNVPGSSLVRESDLAMTNAGTEIGVASTKAPT	407		
Db	361	SQSGETADTIALRLAKTQGYLCAIACNGARSTLVRESDAVLLTHAGSEIGVASTKST	420		
QY	408	TQTLVLLMLVAKLSRLKG-LDASIEHDIVHGLQALPSRIEQMLSDQKRIEALAEAFSDKH	466		
Db	421	TQLVCLLVTRMIAQAKKILTOEPEDALSAALQRLPDVHEVHLECEADVRCARHFVHAQ	480		
QY	467	HALFLSRGDQVPALLEGALKLKEISYTHAEAYAGELKHGPLALIDADMPVIVVAPNNB-	525		
Db	481	HALFUGRGELPYIATESALKLKEISYTHAEAYAGELKHGPLALVDAQMPVVAIPASPG	540		
QY	526	-LLEKLKSNIEVRARGGQLYVFAD-----QDAGFVSSDNHIIEMPHVEVIAPI	574		
Db	541	VLFEKASNIEVRARGGWLXIFTDVPFRFGVCTPEADAPGACSQIVTVPSVPLTAP	600		
QY	575	IFYTVPLQLLAYHVALIKGTDVQDPRNLAKSVTVE 609			
Db	601	IFYAVPLQLLAYHIALKGTIDQDPRNLAKSVTVE 635			

RESULT 11  
A82844  
Glutamine-fructose-6-phosphate transaminase (isomerizing) (EC 2.6.1.16) [similarity] - X  
C:Species: Xylella fastidiosa  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004  
C:Accession: A82844  
R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen  
Nature 406, 151-157, 2000  
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A:Reference number: A82515; MUID:20365717; PMID:10910347  
A:Note: for a complete list of authors see reference number A59328 below  
A:Accession: A82844  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-609 <SIM>  
A:Cross-references: UNIPROT:Q9PH05; UNIPARC:UPI0000165A4A; GB:AE003867; GB:AE003849; NID:  
A:Experimental source: strain 95C  
R:Simpson, A. J. G.; Reinach, F. C.; Arruda, P.; Abreu, F. A.; Acencio, M.; Alvarenga, R.; A  
Briones, M. R. S.; Bueno, M. R. P.; Camargo, A. A.; Camargo, L. E. A.; Carraro, D. M.; Carrer, H.  
as-Neto, E.; Docena, C.; El-Dorriy, H.; Facincanl, A. P.; Ferreira, A. J. S.  
Submitted to GenBank, June 2000  
A:Authors: Ferreira, V. C. A.; Ferro, J. A.; Fraga, J. S.; Franca, S. C.; Franco, M. C.; Frohme  
J. D.; Junqueira, M. L.; Kemper, E. L.; Kitajima, J. P.; Krieger, J. E.; Kuranae, E. E.; Laigre  
chado, M. A.; Madeira, A. M. B. N.; Madeira, H. M. F.; Marino, C. L.; Marques, M. V.; Martins, E.

A;Authors: Martins, E.M.P.; Matsukuma, A.Y.; Menck, C.P.M.; Miracca, E.C.; Miyaki, C.Y.; F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.; Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z A;Reference number: A59328  
A;Contents: annotation  
C;Genetics:  
A;Gene: XF0141  
C;Superfamily: glutamine-fructose-6-phosphate aminotransferase (isomerizing)  
C;Keywords: aminotransferase; isomerase  
P;2-609/Product: glutamine-fructose-6-phosphate transaminase (isomerizing) #status predicted  
F;2/Active site: Cys #status predicted

Query Match	54.4%;	Score 1681;	DB 2;	Length 609;
Best Local Similarity	54.6%;	Pred. No. 4.2e-101;		
Matches	333;	Conservative 103;	Mismatches 172;	Indels 2;
Gaps	2;			

Qy 1 MCGIVGAIAORDVAEIILEGRLRLEYRGYSAGLAVVDTEGHMTRRLRLRGKVMQAQAAE 60  
Db 1 MCGIVGAIAGRDVPVLEGLKRLLEYRGYDSSGLAVLES-GSIRRVRTGEVAMVAAT 59  
Qy 61 EHPHGGTGTIAHTRWATHGSPSEVNAHPHVS-EHIVVHNGIIEHHEPLREBELKARGYTFV 120  
Db 60 QEGFTASLGIGHTRWATHGVTTEANAHPHVSHGVVLVHNGIIEHHEVQRRLSALGYVFQ 119  
Qy 121 SETDTEVIAHLNVNELKOGGTLREAVLRAIPOLRGAYGTVMDSRHPDPTLLAARSGSPV 180  
Db 120 SQDTEVIAHLIHYHMQGDDLLGALCAVKALGIYALAVMSAEPRFCARMGCPFL 179  
Qy 181 IGLMGFNFIASDQALLPVRTRFIFLEEGDIAEITRRSVNIPOKTAEVKRODIESNLQ 240  
Db 180 IGIGDGEHLVASDISAVIQATROVIFLEDGDTAIRDGISIFNAEOCPVERPLHLSNV 239  
Qy 241 YDAGDKGIYCHYMOKEIEYOPNAIKNTLTGRISHGQVDLSELGNADLLESKVEHIQILA 300  
Db 240 LSSLELGEFRHFMQKEIHEQPRVLADTWEAIDAAGPPMLFGAQAESVFRGITGIQILA 299  
Qy 301 CGTSYNSGMVSRWFESLAGIPCDVEITASEFRYKSAVRNSLMITLSOSGETADTLA 360  
Db 300 CGTSYAGLARYWIEAIGLPCHEVASEYRYKAYVNPQHVLVTTISOSGETLDTLEAL 359  
Qy 361 RLSELKELGYLSAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFPTTQTLVLLMLVAKL 420  
Db 360 KYAKALGHRHTLSICNAPDAIPRISBLICYTRAGPEIGVASTKAFPTTQTLVLLMLVAKL 419  
Qy 421 SRLKG-LDASIEHDIVHGLQALPSRIEQLMSQDKRIEALAEFDFSKHHLFLSRGDQYPI 479  
Db 420 GVLRGAVDAEHAAYLEQLRQPCGVQOALNLEPQIAAWEAFASRHHALFLGRGLHVP 479  
Qy 480 ALEGALKKEISYTHAEAYAGELKHGPLALIDADMPVIVVAPNNELLEKLSNIEVRA 539  
Db 480 ALEGALKKEISYTHAEAYAGELKHGPLALVDADMPVIVVAPNDLSLEKYSNMQSVRA 539  
Qy 540 RGGOLYVFADQDAGFVSDNNHIIEMPHVEEVIPIETVPLQLLAVHVALIKGTDDVQ 599  
Db 540 RGSLELFVADQDSHFSEGLHVIRTLRHGTGLVSLPVLHTIPVQLLAYHTALVRGTDVQ 599  
Qy 600 RNLAKSVTVE 609  
Db 600 RNLAKSVTVE 609

RESULT 12  
T45493  
glutamine-fructose-6-phosphate transaminase (isomerizing) (EC 2.6.1.16) glms [similarity  
N;Alternate names: glucosamine-6-phosphate synthase  
C;Species: Thibacillus ferrooxidans  
C;Date: 31-Jan-2000 #sequence\_revision 31-Jan-2000 #text\_change 09-Jul-2004  
C;Accession: T45493  
R;Oppon, J.C.; Sarnovsky, R.J.; Craig, N.L.; Rawlings, D.E.  
J. Bacteriol. 180, 3007-3012, 1998  
A;Title: A Tn7-like transposon is present in the glms region of the obligately chemoaut

A;Reference number: Z22992; MUID:98269023; PMID:9603897  
A;Accession: T45493  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-611 <OPP>  
A;Cross-references: UNIPROT:Q56275; UNIPARC:UPI000016E22C; EMBL:AF032884; NID:g2653994;  
A;Experimental source: ATCC 33020  
C;Genetics:  
A;Gene: glms  
C;Superfamily: glutamine-fructose-6-phosphate aminotransferase (isomerizing)  
C;Keywords: aminotransferase; isomerase

Query Match	51.1%;	Score 1579;	DB 2;	Length 611;
Best Local Similarity	52.4%;	Pred. No. 1.7e-94;		
Matches	322;	Conservative 101;	Mismatches 182;	Indels 10;
Gaps	5;			

Qy 1 MCGIVGAIAORDVAEIILEGRLRLEYRGYSAGLAVVDTEGHMTRRLRLRGKVMQAQAAE 60  
Db 1 MCGIVGVSKTDLVPMILEGLQRLLEYRGYSAGLAILGADADLLRVRSGRVAELTAADV 60  
Qy 61 EHPHGGTGTIAHTRWATHGSPSEVNAHPHVS-EHIVVHNGIIEHHEPLREBELKARGYTF 119  
Db 61 ERGLQGVGIGHTRWATHGVRNCPHMSHEQIAVHNGIIEHFAHLRAHLSAAGYTF 120  
Qy 120 VSETDTEVIAHLNVNELKOGGTLREAVLRAIPOLRGAYGTVMDSRHPDPTLLAARSGSPL 179  
Db 121 TSETDTEVIAHLVHHYRQTAPDLFAATRRAGDLRGAYATAVSSGDPETVCVARMGCP 180  
Qy 180 VIGLGMGFNIASDQALLPVRTRFIFLEEGDIAEITRRSVNIPOKTAEVKRODIESNL 239  
Db 181 LLGVADDGHYFASDVAALLPVRTRVLYLEDGDVAMLQRTLRITDQAGASRQREHWSQL 240  
Qy 240 QYDAGDKGIYCHYMOKEIEYOPNAIKNTLTGRISHGQVDLSEL-GPNADLLESKVEHIQ 298  
Db 241 SAAAVDLGPYRHFQKEIHEQPRAVADTLEGAL-NSQLDLDLMDGGAAMFRDVRDLF 299  
Qy 299 LACTSYNSGMVSRWFESLAGIPCDVEITASEFRYKSAVRNSLMITLSOSGETADTLA 358  
Db 300 LAGSTHYATLVGQWVESIVGIPQAELGHEYRYRDSIPDPQOLVVTLSOSGETLDTFE 359  
Qy 359 GLRLSELKELGYLSAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFPTTQTLVLLMLVA 418  
Db 360 ALRRKDLGHTRTLAI CNVAESAIPRASALRFLTRAGPEIGVASTKAFPTTQLAALYLAL 419  
Qy 419 KLSRLKGLDASIEHDIVHG----LOALPSRIEQLMSQDKRIEALAEFDFSKHHLFLSRG 474  
Db 420 SLAKAPG---ASERCAAGSGPGRRLQPLPGSVQHALNLEPQIQGWAARFASKDHFLGRG 476  
Qy 475 DQYPIALEGALKKEISYTHAEAYAGELKHGPLALIDADMPVIVVAPNNELLEKLSNI 534  
Db 477 LHYPPIALEGALKKEISYTHAEAYAGELKHGPLALVDADMPVIVVAPNDELLEKLANM 536  
Qy 535 EYVRARGOLYVFADQDAGFVSDNNHIIEMPHVEEVIPIETVPLQLLAVHVALIKGT 594  
Db 537 QEVHARGELLYVFADSDSHFNAGVHVMRLPRHAGLLSPVHAIPVQLLAYHAALVKGT 596  
Qy 595 DVDQPRNLAKSVTVE 609  
Db 597 DVDPRNLAKSVTVE 611

RESULT 13  
H82022  
glutamine-fructose-6-phosphate transaminase (isomerizing) (EC 2.6.1.16) NMA0276 [similar  
N;Alternate names: glucosamine fructose-6-phosphate aminotransferase [misnomer]  
C;Species: Neisseria meningitidis  
C;Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 09-Jul-2004  
C;Accession: H82022  
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel  
; Holroyd, S.; Jagsels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,  
Nature 404, 502-506, 2000  
A;Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.  
A;Reference number: A81775; MUID:20222556; PMID:10761919



A;Accession: H82022  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-612 <PAR>  
A;Cross-references: UNIPROT:Q9UWN9; UNIPARC:UPI0000165682; GB:AL162752; GB:AL157959; NID  
A;Experimental source: serogroup A, strain Z2491  
C;Genetics:  
A;Gene: glmS; NMA0276  
C;Superfamily: Glutamine-fructose-6-phosphate aminotransferase (isomerizing)  
C;Keywords: aminotransferase; isomerase  
F;2-612/Product: glutamine-fructose-6-phosphate transaminase (isomerizing) #status predicted  
F;2/Active site: Cys #status predicted

Query Match	50.8%;	Score 1569.5;	DB 2;	Length 612;
Best Local Similarity	52.8%;	Pred. No. 7.1e-94;		
Matches 325;	Conservative 104;	Mismatches 177;	Indels 9;	Gaps 9;
QY	1	MCQIVGAI-AOROVAETLLEGRLEVRGYDSAGLAVVDTEGHMTLRRLRGKQVQLMAQAA	59	
DB	1	MCQIVGAI-RAHNVVDFLTDLGRLEVRGYDSSGIA-VNTDGKIKRVRRVGRVQLMEDAA	59	
QY	60	EEHPLHGCTGIAHTRWATHGEPSEVNAHPVS-EHIVVVHNGIIENHEPRLRELKARGYT	118	
DB	60	REXGISGGIGIGHTRWATHGCVTEPNAPHHISGMIADVHNGIIENFESERKLEGLGYR	119	
QY	119	FVSETTEVTAHLAVNNELKQ-GGTLREAVLRAIPOLRGAYGTVIMDSRHPDPTLLAARSGS	177	
DB	120	FESQTDTEVTAHSINHEYAONGKLFPAVQEAVKRTHGAYAIATAQDRPELVVARMGC	179	
QY	178	PLVIGLGMGNFIASDQALLPVTTRPIFLEEEDGDIABITRRSV-NIFDKTGAEVKRDQIE	236	
DB	180	PLLVALGDDETFIASDVSAVIAFTRRVAYLEDGDIALLASDGIKRLTKDSGLPAERKVKV	239	
QY	237	SNLOYDAGDKGIYCHYWOKEIYEQPNAIKNTLGTGRISHGQVDLSELGPNADELLSKVEHI	296	
DB	240	SELSLASLELPGYSHFMQKEIHEQOPRAIADTAEVFLDGGFIP-ENFGKNAKSVFESIRSV	298	
QY	297	QILACGTSYNGSMVSRWPFESLAGIPCDBEIASFEFRRYKSAVRNRNSIMITLSQSGETADT	356	
DB	299	KILACGTSYAAALTAKVWLESIAKIPSDVEIASERYRSVIADPDQLVITISQSGETLDT	358	
QY	357	LAGIRLSKEIYGLGSLAICNVPGSSLVRESDLALMNTNAGTEIGVASTKATPTTQTLVLML	416	
DB	359	MEALKYAKSLGHRHLSLINCNMESALPRESLSVLYTRAGAEIGVASTKATPTTQTLVALFGL	418	
QY	417	VAKLSRLKGL-DASIEHDI VHGLQALPSRTEOMLSODKRIEALAEFSDKHAFALFSRGD	475	
DB	419	AVTLAKVRGLVSEDEARYTEELRQLPGSVQHALLNLEPQIAANAQQFAKTSALFLGRGI	478	
QY	476	QYPIALEGALKLKEISYTHAEVAAAGELKHGPIALADAMPVIVAPNNELLBKLSKNIE	535	
DB	479	HYPIALEGALKLKEITYTHAEVAPAGELKHGPIALVDENMPVVIATPNDSLDDKVKANMQ	538	
QY	536	EVARGQLVVFADQDAGFVSSNMHIEMP-HVEEVIAPIFYTVPLQILLAYHVALIKGT	594	
DB	539	EVARGELLEVFADLDSNFNATEGVHVIRAPRVGE-LSPVVTHTPVQLLSYHVALARGT	597	
QY	595	DVDQPRNLAKSVTVE	609	
DB	598	DVDQPRNLAKSVTVE	612	

RESULT 14  
B81246  
glutamine-fructose-6-phosphate transaminase (isomerizing) (EC 2.6.1.16) NMB0031 [similar  
N; alternate names: glucosamine fructose-6-phosphate aminotransferase [misnomer]  
C; Species: *Neisseria meningitidis*  
C; Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 09-Jul-2004  
C; Accession: B81246  
R; Rettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.;  
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;  
Ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.  
Science 287, 1809-1815, 2000

A; Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ver  
A; Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58.  
A; Reference number: AB1000; MUID:2015755; PMID:10710307  
A; Accession: B81246  
A; Status: preliminary  
A; Molecule type: DNA  
A; Residues: 1-612 <TEST>  
A; Cross-references: UNIPROT:Q9K1P9; UNIPARC:UPI000016563D; GB:AE002361; GB:AE002098; NID:  
A; Experimental source: serogroup B, strain MC58

Query Match	50.7%;	Score 1565.5;	DB 2;	Length 612;	
Best Local Similarity	52.3%;	Pred. No. 1.3e-93;			
Matches 321;	Conservative 104;	Mismatches 182;	Indels 7;	Gaps 7;	

  

QY	1	MC	G	I	G	A	I	-	A	O	R	D	I	A	E	T	L	L	E	G	L	R	E	Y	G	D	S	A	G	L	V	V	D	T	E	G	H	M	T	R	L	R	L	G	K	V	M	L	A	O	A	59										
DB	1	MC	G	I	G	A	I	-	A	O	R	D	I	A	E	T	L	L	E	G	L	R	E	Y	G	D	S	S	G	I	A	-	V	N	T	D	G	K	I	K	R	V	R	G	E	V	L	M	E	D	A	59										
QY	60	E	E	H	L	H	G	G	T	G	I	A	T	R	W	A	T	H	G	E	P	S	E	V	N	A	H	P	H	S	-	E	H	I	V	V	H	N	G	I	I	E	N	H	E	P	L	R	E	E	L	K	A	R	G	T	118					
DB	60	R	E	X	I	S	G	I	G	I	G	I	T	R	W	A	T	H	G	E	P	N	A	H	P	I	S	G	M	I	A	V	V	H	N	G	I	I	E	N	F	E	S	E	R	K	R	E	L	G	Y	119										
QY	119	F	V	S	E	T	T	E	V	I	A	H	L	V	N	N	E	L	K	O	-	G	T	L	R	E	A	V	L	R	A	I	P	O	L	R	G	A	T	V	I	M	D	S	R	H	P	T	L	L	A	A	R	S	G	177						
DB	120	F	S	E	Q	T	D	E	V	I	A	H	S	I	N	H	E	Y	A	O	G	R	L	F	E	A	V	Q	E	A	V	R	F	H	G	A	I	A	V	I	A	O	K	P	D	E	L	V	A	R	M	C	179									
QY	178	P	L	V	I	G	L	M	G	E	N	F	I	A	S	D	O	L	L	P	T	R	R	F	I	F	L	E	E	G	D	I	A	E	I	T	R	S	V	-	N	I	F	D	K	T	G	A	E	V	K	R	D	I	E	236						
DB	180	P	L	L	V	A	L	G	D	D	E	T	F	I	A	S	D	S	A	V	I	A	T	P	R	R	V	A	L	E	D	G	D	I	A	L	L	A	S	D	G	I	K	R	L	T	D	K	N	G	L	P	A	E	R	K	V	239				
QY	237	S	N	L	O	Y	D	A	G	D	K	G	I	I	C	H	Y	N	O	K	E	I	Y	E	O	P	N	A	I	K	N	T	L	T	G	R	I	S	H	G	O	V	D	L	S	E	L	G	P	N	A	D	E	L	L	S	K	V	E	H	I	296
DB	240	S	E	L	S	A	S	L	E	G	L	Y	S	H	F	O	K	E	I	H	E	O	P	R	A	I	A	D	T	A	E	V	L	D	G	G	F	I	P	-	E	N	F	G	K	D	A	K	S	V	F	E	S	I	R	V	298					
QY	297	Q	I	L	A	C	G	T	S	Y	N	G	M	V	R	S	Y	F	E	S	L	A	G	I	P	C	D	O	B	I	A	S	E	F	R	Y	K	R	S	A	V	R	R	N	S	L	M	I	T	S	Q	S	E	T	A	D	T	356				
DB	299	K	I	L	A	C	G	T	S	Y	A	A	L	T	A	K	Y	L	E	S	I	A	K	I	P	S	D	V	E	I	A	S	E	V	R	S	V	I	A	D	S	D	L	Q	L	V	I	T	S	Q	S	E	T	I	D	T	358					
QY	357	L	A	G	I	R	S	K	E	L	G	Y	G	S	L	A	I	C	N	V	P	S	S	L	V	R	E	S	D	L	A	M	T	N	A	G	T	E	I	G	V	A	S	T	K	A	F	T	L	T	V	L	M	L	416							
DB	359	M	E																																																											

RESULT 15  
AD3595  
glutamine-fructose-6-phosphate transaminase (isomerizing) (EC 2.6.1.16) [imported] - Bruc  
C:Species: Brucella melitensis  
C:Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 09-Jul-2004  
C:Accession: AD3595  
R:BelVecchio, V. G.; Kapatal, V.; Redkar, R. J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, N  
; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P. H.; Hagius, S.; O'Callaghan, D.; Letessier  
C:Accession: AD3595



Search completed: June 14, 2006, 15:28:37  
Job time : 43.2344 secs

**THIS PAGE IS BLANK**

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 14, 2006, 15:14:44 ; Search time 74.7383 Seconds

(without alignments)  
972.693 Million cell updates/sec

Title: US-10-612-779-30

Perfect score: 835

Sequence: 1 MSLLPGFFVIRNMEEGDLEQV.....VKFYKGFNSNAGVEMQIRK 159

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_8.\*

- 1: geneseqp1380s.\*
- 2: geneseqp1390s.\*
- 3: geneseqp2000s.\*
- 4: geneseqp2001s.\*
- 5: geneseqp2002s.\*
- 6: geneseqp2003as.\*
- 7: geneseqp2003bs.\*
- 8: geneseqp2004s.\*
- 9: geneseqp2005s.\*
- 10: geneseqp2006s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	835	100.0	159	8	ADI38885 Glucosami
2	835	100.0	159	8	ADJ64176 S. cerevi
3	835	100.0	159	8	ADU00600 Amino aci
4	333	39.9	149	5	ABP73464 Candida a
5	328	39.3	149	8	ADI38887 Glucosami
6	328	39.3	149	8	ADU00602 Amino aci
7	275	32.9	190	6	ABJ26242 Aspergill
8	261.5	31.3	149	8	ADI38889 Glucosami
9	216.5	25.9	174	8	ADY05609 Plant ful
10	209	25.0	184	4	AAU07779 Human nov
11	208	24.9	184	4	AAU07779 Human tra
12	208	24.9	184	4	AAU07779 Human tra
13	208	24.9	184	8	ADRI0847 Human FLJ
14	208	24.9	184	8	ADRI0848 Human FLJ
15	208	24.9	200	4	AAU07779 Human nov
16	207	24.8	219	4	ABBS8239 Drosophil
17	204	24.4	184	5	ABU65053 Human NOV
18	203	24.3	206	4	AAU79992 Human pro
19	202.5	24.3	180	6	ABJ25642 Aspergill
20	202	24.2	184	5	ABU65054 Human NOV
21	202	24.2	184	8	ADN61759 Human nov
22	199	23.8	184	4	AAU79008 Human pro
23	189.5	22.7	183	8	ADN61757 Human nov

24	158	18.9	107	4	AAU23427	Novel hum
25	126	15.1	154	5	ABP27501	Streptoco
26	126	15.1	154	8	ADV81684	Streptoco
27	126	15.1	156	8	ADV88261	Streptoco
28	126	15.1	156	8	ADV79514	Streptoco
29	117.5	14.1	242	7	ABO79472	Pseudomon
30	116.5	14.0	157	5	ABB50043	Listeria
31	113	13.5	154	5	ABB53909	Lactococc
32	112	13.4	140	6	ABU23792	Protein e
33	111	13.3	166	6	ABW72742	Staphyloc
34	110	13.2	179	4	AAU78880	C. glutam
35	110	13.2	179	4	AAU90861	C. glutam
36	106	12.7	149	7	ADC95669	E. faeciu
37	106	12.7	185	5	ABB47559	Listeria
38	106	12.7	306	6	ABW70329	Photorhab
39	105.5	12.6	159	5	AAE25024	Human dru
40	104	12.5	307	7	ADF06564	Bacterial
41	102	12.2	154	9	ABE42282	L. pneumo
42	102	12.2	184	9	ABE36279	L. pneumo
43	101	12.1	134	8	ADP30044	Human sec
44	101	12.1	168	5	ABB49030	Listeria
45	101	12.1	168	6	ABU32907	Protein e

ALIGNMENTS

RESULT 1

ADI38885

ID ADI38885 standard; protein; 159 AA.

XX AC ADI38885;

DT 15-APR-2004 (first entry)

DE Glucosamine-6-phosphate acetyltransferase, GNA1, SEQ ID 30.

KW Glucosamine; N-acetylglucosamine; fermentation;

KW glucosamine-6-phosphate acetyltransferase;

KW glucosamine-6-phosphate synthase; glucosamine-6-phosphate deaminase;

KW glucosamine-1-phosphate N-acetyltransferase; glucosamine-6-phosphate;

KW glucosamine-1-phosphate; N-acetylglucosamine-1-phosphate;

KW N-acetylglucosamine-6-phosphate; enzyme.

XX OS Saccharomyces cerevisiae.

XX WO2004003175-A2.

XX PD 08-JAN-2004.

XX PF 01-JUL-2003; 2003WO-US020925.

XX PR 01-JUL-2002; 2002US-0393348P.

XX PA (ARKI-) ARKION LIFE SCI LLC.

PI Deng M, Angerer JD, Cyron D, Grund AD, Jerrell TA, Leanna C;

PI Mathre O, Rosson R, Running J, Severson D, Song L, Wassink S;

XX WPI; 2004-203380/19.

XX DR N-PSDB; ADI38884.

XX PT Producing glucosamine or N-acetylglucosamine by fermentation involves

PT culturing microorganism comprising glucosamine-6-phosphate

PT acetyltransferase, in fermentation medium, and collecting product.

XX PS Claim 5; SEQ ID NO 30; 327pp; English.

XX CC The present invention relates to a method (M1) for producing glucosamine

XX CC and N-acetylglucosamine by fermentation. The method comprises (a)

XX CC culturing in a fermentation medium a microorganism (I) which comprises

XX CC endogenous glucosamine-6-phosphate acetyltransferase (II) and a genetic

XX CC modification that increases the activity of (II), glucosamine-6-phosphate

CC synthase (III) or glucosamine-6-phosphate deaminase (IV), or decreases  
CC the activity of (IV) and increases the activity of glucosamine-1  
CC phosphate N-acetyltransferase (V), and (b) and collecting the product,  
CC which is chosen from the group consisting of glucosamine-6-phosphate, N-  
CC glucosamine, glucosamine-1-phosphate, N-acetylglucosamine-1-phosphate, N-  
CC acetylglucosamine-6-phosphate, and N-acetylglucosamine. The present  
CC sequence was used to illustrate the method of the invention.  
XX  
SQ Sequence 159 AA;  
  
Query Match 100.0%; Score 835; DB 8; Length 159;  
Best Local Similarity 100.0%; Pred. No. 3e-83;  
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 MSLPDGFYIRMEEGDLEQVETTLKVLTTVGTITPESFSKLIKYWNEATVWVNDNEDKKIM 60  
Db |||||||  
1 MSLPDGFYIRMEEGDLEQVETTLKVLTTVGTITPESFSKLIKYWNEATVWVNDNEDKKIM 60  
Qy 61 QYNPMVIVDKRTETVAATGNIIIERKIIHELGLCGHIEDIAVNSKYQGGLGKLLIDQLV 120  
Db |||||||  
61 QYNPMVIVDKRTETVAATGNIIIERKIIHELGLCGHIEDIAVNSKYQGGLGKLLIDQLV 120  
  
Qy 121 TIGFDYGCYKIILDCDEKNVFKYKCGFSNAGVEMQIRK 159  
Db |||||||  
121 TIGFDYGCYKIILDCDEKNVFKYKCGFSNAGVEMQIRK 159  
  
RESULT 2  
ADJ64176  
ID ADJ64176 standard; protein; 159 AA.  
AC ADJ64176;  
DT 06-MAY-2004 (first entry)  
DE S. cerevisiae glucosamine phosphate N-acetyltransferase.  
KW Hexosamine; arthritis; osteoarthritis; rheumatoid arthritis; pain;  
KW inflammation; articular joint disease; spondyloarthritis;  
KW gouty arthritis; systemic lupus erythematosus; juvenile arthritis;  
KW tendinitis; bursitis; connective tissue injury; psoriasis; eczema; burn;  
KW dermatitis; transgenic plant; glucosamine phosphate N-acetyltransferase;  
KW yeast; enzyme; EC 2.3.1.4.  
XX  
OS Saccharomyces cerevisiae.  
XX  
PN US2004003432-A1.  
XX  
PD 01-JAN-2004.  
XX  
PF 05-MAY-2003; 2003US-00429812.  
XX  
PR 06-MAY-2002; 2002US-0378297P.  
XX  
PA (PHAA ) PHARMACIA CORP.  
XX  
PI Obukowicz MG;  
XX  
DR WPI; 2004-061704/06.  
DR N-PSDB; ADJ64175.  
XX  
PT Producing hexosamine useful for treating arthritis by culturing cell  
PT comprising heterologous polynucleotide sequences which encode for enzymes  
PT required for biosynthetic pathway capable of synthesizing hexosamine.  
XX  
PS Disclosure; SEQ ID NO 4; 35pp; English.  
XX  
CC The invention relates to a method of producing a hexosamine that involves  
CC providing a cell comprising polynucleotide sequences which code for  
CC enzymes required for a biosynthetic pathway capable of synthesizing the  
CC hexosamine, where at least one of the polynucleotide sequences comprises  
CC a recombinant polynucleotide. The method involves transforming a cell  
CC with at least one heterologous polynucleotide coding for a polypeptide in

CC a biosynthetic pathway that is capable of producing hexosamine, culturing  
CC the transformed cell under conditions that permit the cell to translate  
CC the polynucleotide into a polypeptide comprising an enzyme which is part  
CC of the biosynthetic pathway. The hexosamine produced by the method of  
CC invention is useful for treating arthritis, preferably osteoarthritis and  
CC rheumatoid arthritis in humans and pet animals. The hexosamines are also  
CC useful for treating pain, inflammation or inflammation-associated  
CC disorder. The hexosamines are also useful for treating other articular  
CC joint damage or diseases such as spondyloarthritis, gouty arthritis,  
CC systemic lupus erythematosus and juvenile arthritis and other joint  
CC damage disease conditions such as tendinitis, bursitis, connective tissue  
CC injuries or disorders, and skin related conditions such as psoriasis,  
CC eczema, burns and dermatitis. The method provides improved production of  
CC hexosamine without compromising its clinical effectiveness and in a form  
CC that is acceptable for delivery to a broad class of patients. Consumption  
CC of edible portions of transgenic plants containing hexosamines provide a  
CC supply of hexosamine to the subject in a conventional and easily  
CC consumable form. The present sequence is Saccharomyces cerevisiae  
CC glucosamine phosphate N-acetyltransferase (EC 2.3.1.4).  
XX  
SQ Sequence 159 AA;

Query Match 100.0%; Score 835; DB 8; Length 159;  
Best Local Similarity 100.0%; Pred. No. 3e-83;  
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 MSLPDGFYIRMEEGDLEQVETTLKVLTTVGTITPESFSKLIKYWNEATVWVNDNEDKKIM 60  
Db |||||||  
1 MSLPDGFYIRMEEGDLEQVETTLKVLTTVGTITPESFSKLIKYWNEATVWVNDNEDKKIM 60  
Qy 61 QYNPMVIVDKRTETVAATGNIIIERKIIHELGLCGHIEDIAVNSKYQGGLGKLLIDQLV 120  
Db |||||||  
61 QYNPMVIVDKRTETVAATGNIIIERKIIHELGLCGHIEDIAVNSKYQGGLGKLLIDQLV 120  
  
Qy 121 TIGFDYGCYKIILDCDEKNVFKYKCGFSNAGVEMQIRK 159  
Db |||||||  
121 TIGFDYGCYKIILDCDEKNVFKYKCGFSNAGVEMQIRK 159

RESULT 3  
ADU00600  
ID ADU00600 standard; protein; 159 AA.  
XX  
AC ADU00600;  
XX  
DT 13-JAN-2005 (first entry)  
XX  
DE Amino acid sequence of glucosamine-6-phosphate acetyltransferase.  
XX  
KW chitin; chitosan; fermentation;  
KW glutamine-fructose-6-phosphate amidotransferase; glucosamine synthetase;  
KW glucosamine-6-phosphate synthase; GFA1 gene; glms gene;  
KW glucosamine-6-phosphate acetyltransferase; GNA1 gene; chitin synthase;  
KW chitin deacetylase; CDA1 gene; CDA2 gene;  
KW N-acetylglucosamine-6-phosphate deacetylase;  
KW glucosamine-6-phosphate deaminase; chitinase; chitosanase; fungus; yeast;  
XX enzyme.  
XX  
OS Saccharomyces cerevisiae.  
XX  
PN WO2004092391-A2.  
XX  
PD 28-OCT-2004.  
XX  
PF 12-APR-2004; 2004WO-US011286.  
XX  
PR 11-APR-2003; 2003US-0462087P.  
XX  
PA (ARKI-) ARKION LIFE SCI LLC.  
XX  
PI Deng M, McMullin TW, Grund AD;  
XX WPI; 2004-766880/75.  
DR

DR N-PSDB; ADU00599.  
 XX Producing chitin or chitosan, involves culturing microorganism with  
 PT genetic modification that results in increase in activity of glutamine-  
 PT fructos-6-phosphate amidotransferase, in fermentation medium, to produce  
 PT chitin or chitosan.  
 XX  
 XX Example 1; SEQ ID NO 33; 161pp; English.  
 XX  
 XX The specification describes a method for producing chitin or chitosan by  
 CC a fermentation process. The method involves culturing in a fermentation  
 CC medium a microorganism which comprises one or more genetic modifications  
 CC that result in an increase in the activity of glutamine-fructose-6-  
 CC phosphate amidotransferase (also known as glucosamine synthetase and  
 CC glucosamine-6-phosphate synthase), and encoded by the GNA1 eukaryotic gene  
 CC and the glnS bacterial gene), glucosamine-6-phosphate acetyltransferase  
 CC (encoded by GNA1), chitin synthase or chitin deacetylase (encoded by CDA1  
 CC and CDA2), or in decrease in the activity of N-acetylglucosamine-6-  
 CC phosphate deacetylase, glucosamine-6-phosphate deaminase, chitinase and  
 CC chitosanase, and collecting the chitin or chitosan. The method is useful  
 CC for producing chitin or chitosan by utilizing microorganisms such as  
 CC fungus, yeast (e.g. Saccharomyces or Schizosaccharomyces) and  
 CC filamentous fungus (e.g. Aspergillus, Absidia or Rhizopus), preferably S.  
 CC cerevisiae, A. niger or A. nidulans. The method enables high quantities  
 CC of chitin and chitosan to be produced cost effectively. The present  
 CC sequence is encoded by a GNA1 gene, and is a glucosamine-6-phosphate  
 CC acetyltransferase. The GNA1 gene was used to transform yeast for use in  
 CC the method of the invention.  
 XX  
 XX Sequence 159 AA;

Query Match 100.0%; Score 835; DB 8; Length 159;  
 Best Local Similarity 100.0%; Pred. No. 3e-83;  
 Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MSLPDGFYIRMEEGDLEQVTTTLKVLTVGTITPESFSLIKYWNATVWMDNEDKKIM 60  
 Db 1 MSLPDGFYIRMEEGDLEQVTTTLKVLTVGTITPESFSLIKYWNATVWMDNEDKKIM 60  
 QY 61 QYNPMVVDKRTETVAATGNIIIRKIIHELGLCGHIEDIAVNSKYQGGLKLLIDQLV 120  
 Db 61 QYNPMVVDKRTETVAATGNIIIRKIIHELGLCGHIEDIAVNSKYQGGLKLLIDQLV 120  
 QY 121 TTGFDYGCYKIIILDCDEKNVCFYKCGFSNAGVEMQIRK 159  
 Db 121 TTGFDYGCYKIIILDCDEKNVCFYKCGFSNAGVEMQIRK 159

RESULT 4  
 ABP73464  
 ID ABP73464 standard; protein; 149 AA.

AC ABP73464;  
 XX 30-JAN-2003 (first entry)  
 DT  
 XX  
 XX Candida albicans essential protein SEQ ID NO 7301.  
 DE  
 XX Fungus; yeast; tetracycline; promoter; GRACE strain; biosynthesis;  
 KW signal transduction; DNA replication; cell division; growth;  
 KW proliferation; Candida albicans; fungicide; antifungal.  
 XX  
 XX Candida albicans.  
 OS  
 XX WO200253728-A2.  
 PN  
 XX 11-JUL-2002.  
 PD  
 XX 26-DEC-2001; 2001WO-US049486.  
 PF  
 XX 29-DEC-2000; 2000US-0259128P.  
 XX  
 PR 20-FEB-2001; 2001US-00792024.  
 PR 22-AUG-2001; 2001US-0314050P.

XX (ELIT-) ELITRA PHARM INC.  
 XX Roemer T, Jiang B, Boone C, Bussey H, Ohlsen KL;  
 PI WPI; 2002-566694/60.  
 XX N-PSDB; ABZ32014.  
 DR  
 XX Constructing strains for identifying gene products as effective targets  
 PT for therapeutic intervention, by inactivating in the strain one allele of  
 PT a gene and placing other allele of the gene under conditional expression.  
 PT  
 XX Claim 44; SEQ ID NO 7301; 167pp + Sequence Listing; English.  
 XX  
 XX The invention relates to constructing (M1) a strain of diploid fungal  
 CC cells in which both alleles of a gene are modified, comprising modifying  
 CC one allele by insertion or replacement by a cassette having an  
 CC expressible selectable marker and modifying other allele by  
 CC recombination, of a promoter replacement fragment with a heterologous  
 CC promoter, so that expression of the second allele is regulated by the  
 CC promoter. (M1) is useful for constructing a strain of diploid fungal  
 CC cells in which both alleles of a gene are modified. The diploid fungal  
 CC cells having both alleles modified are useful for identifying a gene that  
 CC is essential to the survival or growth of a fungus, a gene that  
 CC contributes to the virulence and/or pathogenicity of a fungus, a gene  
 CC that contributes to the resistance of a diploid fungus to an antifungal  
 CC agent, an antifungal agent that inhibits the growth of a diploid fungus  
 CC and for identifying a therapeutic agent for treatment of a diploid fungus  
 CC disease. (M1) is useful for identifying a compound which modulates the  
 CC activity of a gene product, preferably enzymatic activity, carbon  
 CC compound catabolism, biosynthetic, transporter, transcriptional,  
 CC translational, signal transduction, DNA replication and cell division  
 CC activity. The method is useful for identifying a compound having the  
 CC ability to inhibit growth or proliferation of C. albicans cells and for  
 CC treating infection by C. albicans. The present sequence is that of an  
 CC essential Candida albicans protein used in the method of the invention.  
 CC Note: The sequence data for this patent is not represented in the printed  
 CC specification but is based on sequence information supplied to Derwent by  
 CC the European Patent Office  
 XX

Sequence 149 AA;

Query Match 39.9%; Score 333; DB 5; Length 149;  
 Best Local Similarity 44.0%; Pred. No. 4.2e-28;  
 Matches 70; Conservative 26; Mismatches 51; Indels 12; Gaps 3;  
 QY 1 MSLPDGFYIRMEEGDLE-QVTETLKVLTVTGTTTPESFSLIKYWNATVWMDNEDKKI 59  
 Db 1 MSLPDGFYIRMEEGDLE-QVTETLKVLTVTGTTTPESFSLIKYWNATVWMDNEDKKI 59  
 QY 60 MOYNPMVVDKRTETVAATGNIIIRKIIHELGLCGHIEDIAVNSKYQGGLKLLIDQLV 119  
 Db 51 STYHEPVITN-ASGIIVATGMLFVEKKLIHECGKVGHTIEDISVAKSEQGKGLGYLVTSL 109  
 QY 120 VTIGFDYGCYKIIILDCDEKNVCFYKCGFSNAGVEMQIR 158  
 Db 110 TKVAQENDCYKVLDCSPENGVFYKCGYKDGVEVNCR 148

RESULT 5  
 ADI38887  
 ID ADI38887 standard; protein; 149 AA.

AC ADI38887;  
 XX  
 XX 15-APR-2004 (first entry)  
 DT  
 XX Glucosamine-6-phosphate acetyltransferase, GNA1, SEQ ID 32.  
 DE  
 XX Glucosamine; N-acetylglucosamine; fermentation;  
 KW glucosamine-6-phosphate acetyltransferase;  
 KW glucosamine-6-phosphate synthase; glucosamine-6-phosphate deaminase;  
 KW glucosamine-1 phosphate N-acetyltransferase; glucosamine-6-phosphate;





Dd		47	YGDDHVICVIEETSGKTAATGSMWIEKKFLRNGCKAGHIEDVVDSRFKGQLGKKVVE	108
Qy		118	QLVTIGFDYGCYKIILDCDEKNVRPFYEKCGFSNAGVEM	155
			:       :          :	
Dd		107	FLMDHCKSMGYKVILDGSVENKFYEKCGMSNKSIQM	144
RESULT 9				
ADY05609	standard; protein; 174 AA.			
ID				
AC	ADY05609;			
XX				
DT	21-APR-2005 (first entry)			
XX				
DE	Plant full length insert polypeptide seqid 61424.			
XX				
KW	plant protectant; plant growth regulant; gene therapy; plant;			
KW	recombinant DNA construct; physical array; plant breeding marker;			
KW	cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;			
KW	extreme osmotic condition; pathogen tolerance; pest tolerance;			
KW	growth rate; cell cycle pathway; disease resistance;			
KW	galactomannan production; lignin production; plant growth regulator;			
KW	yield; plant growth; plant development; seed oil; protein yield;			
KW	protein content.			
XX				
OS	Unidentified.			
XX				
PN	US2004034888-A1.			
XX				
PD	19-FEB-2004.			
XX				
PF	28-APR-2003; 2003US-00425114.			
XX				
PR	06-MAY-1999; 99US-00304517.			
PR	05-NOV-2001; 2001US-00985678.			
XX				
PA	(LIU/) LIU J.			
PA	(ZHOU/) ZHOU Y.			
PA	(KOVA/) KOVALIC D K.			
PA	(SCRE/) SCREEN S E.			
PA	(TABAA/) TABASKA J E.			
PA	(CAO/) CAO Y.			
XX				
PI	Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;			
XX				
DR	WPI; 2004-180133/17.			
XX				
PT	New recombinant DNA construct, useful for improving plant tolerance to			
PT	cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or			
PT	pests, for conferring increased resistance to plant disease, or for			
PT	improving yield.			
XX				
PS	Claim 1; SEQ ID NO 61424; 15pp; English.			
XX				
CC	The invention describes a recombinant DNA construct comprising a			
CC	polynucleotide consisting of a sequence encoding an amino acid sequence			
CC	available in electronic form from the US patent office at			
CC	ftp.segdata.uspto.gov/sequence.html?docID:2004034888. The polynucleotide			
CC	of the invention are also useful in physical arrays of molecules and as			
CC	plant breeding markers. The recombinant DNA construct is useful for			
CC	improving plant tolerance to cold, heat, drought, herbicides, extreme			
CC	osmotic conditions, pathogens or pests, for manipulating growth rate in			
CC	plant calls by modification of the cell cycle pathway, for conferring			
CC	increased resistance to plant disease, for producing galactomannan,			
CC	lignin or plant growth regulators, for increasing the rate of homologous			
CC	recombination in plants, for improving yield by modification of			
CC	photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake			
CC	or by providing improved plant growth and development under at least one			
CC	stress condition or for modifying seed oil or protein yield and/or			
CC	content. This is the amino acid sequence of a plant full length insert			
CC	polypeptide that can be used in the recombinant DNA construct of the			
CC	invention.			



Best Local Similarity 32.9%; Pred. No. 2 4e-14;  
Matches 48; Conservative 25; Mismatches 61; Indels 12; Gaps 2;

QY 5 DGFYIRMEEGDLEQ-VTETLKVLTGVTITPESFSLIKYWNATVWMDNEDKIMQYN 63  
Db 37 EGLVLRPLCTADLNRGFFKVLGQLTETGVVSPQFMKSPEHM-----KKSQDY 85

QY 64 PMWIVDKRTETVAATGNIIRKIIHELGLCGHIEDIAVNSKYQGGLGKLLIDQLVTTIG 123  
Db 86 VTVEDVTLGQIVATATLIIHKFIHSCAKRGVEDVVSDCRGKQLKLLSTLTLLS 145

QY 124 FDYGVCIILDCDEKNVPEYKCGFS 149  
Db 146 KKLNCYKITLECLPQNVGFYKVGVT 171

RESULT 11  
AAB73505  
ID AAB73505 standard; protein; 184 AA.  
XX AAB73505;  
XX  
XX  
DT 31-JUL-2001 (first entry)  
XX  
DE Human transferase HTFS-12, SEQ ID NO:12.  
XX  
XX Human transferase; HTFS; agonist; antagonist; cellular signalling;  
KW proliferation; cell proliferative disorder; immune disorder;  
KW atherosclerosis; hepatitis; psoriasis; cancer; tumour; inflammation;  
KW AIDS; Addison's disease; allergy; asthma; anaemia; cirrhosis;  
KW Crohn's disease; atopic dermatitis; diabetes mellitus;  
KW multiple sclerosis; rheumatoid arthritis; pancreatitis;  
KW systemic lupus erythematosus; thrombocytopenia; ulcerative colitis;  
KW haemodialysis; extracorporeal circulation; trauma; transgenic animal;  
KW gene therapy; drug screening.  
XX  
XX Homo sapiens.  
XX  
XX WO200132888-A2.  
XX  
XX 10-MAY-2001.  
XX  
XX 02-NOV-2000; 2000WO-US030485.  
XX  
XX 04-NOV-1999; 99US-0163595P.  
XX  
XX (INCY-) INCYTE GENOMICS INC.  
XX  
XX Tang YT, Yue H, Hillman JL, Lal P, Bandman O, Patterson C;  
PI Shih LL, Azimzai Y, Lu DAM, Baughn MR;  
XX  
XX WPI; 2001-328796/34.  
DR N-PSDB; AAH23812.  
XX  
XX Human transferase polypeptides and polynucleotides useful for diagnosis,  
PT prevention and treatment of cell proliferative and immune system  
PT disorders and for identifying agonists and antagonists.  
XX  
XX Claim 1; Page 111; 157pp; English.  
XX  
XX Sequences AAB73494-AAB73535 represent novel human transferase proteins  
CC HTFS-1 to HTFS-42, and sequences AAH23801-AAH23842 represent cDNAs  
CC encoding them. The proteins play important roles in the regulation of  
CC cellular signalling and proliferation. The HTFS proteins are useful for  
CC screening compounds for their effectiveness as agonists or antagonists of  
CC transferase activity, or for compounds that specifically bind to an HTFS  
CC protein or which modulates the activity of an HTFS protein.  
CC Pharmaceutical compositions comprising an HTFS protein. HTFS agonist or  
CC antagonist, or genetic construct encoding an HTFS protein are useful for  
CC treating a disease or condition associated with decreased or increased  
CC expression of functional HTFS. Disorders which may be treated using such  
CC compositions include cell proliferative disorders and immune disorders.  
CC For example, diseases which may be treated include atherosclerosis,

CC hepatitis, psoriasis, cancers (including breast, bladder, bone marrow,  
CC brain and uterus cancer), inflammation, AIDS, Addison's disease,  
CC allergies, asthma, anaemia, cirrhosis, Crohn's disease, atopic  
CC dermatitis, diabetes mellitus, multiple sclerosis, rheumatoid arthritis,  
CC pancreatitis, systemic lupus erythematosus, thrombocytopenia, and  
CC ulcerative colitis. They may also be used to treat complications of  
CC cancer, haemodialysis, extracorporeal circulation, trauma and  
CC haematopoietic cancer, including lymphoma, leukaemia and myeloma.  
CC Polynucleotides encoding HTFS proteins are useful for creating transgenic  
CC animals to model human diseases, for diagnostic purposes and to generate  
CC hybridisation probes useful in mapping the naturally occurring genomic  
CC sequences. HTFS, and its catalytic or immunogenic fragments are useful  
CC for screening libraries of compounds in a variety of drug screening  
CC techniques. Antibodies which specifically bind HTFS may be used for the  
CC diagnosis of disorders associated with the expression of HTFS, or in  
CC assays to monitor patients being treated with HTFS or agonists,  
CC antagonists or inhibitors of HTFS. The present sequence represents an  
CC HTFS protein of the invention  
XX  
XX Sequence 184 AA;  
SQ

Query Match 24.9%; Score 208; DB 4; Length 184;  
Best Local Similarity 32.9%; Pred. No. 3.1e-14;  
Matches 48; Conservative 25; Mismatches 61; Indels 12; Gaps 2;

QY 5 DGFYIRMEEGDLEQ-VTETLKVLTGVTITPESFSLIKYWNATVWMDNEDKIMQYN 63  
Db 37 EGLVLRPLCTADLNRGFFKVLGQLTETGVVSPQFMKSPEHM-----KKSQDY 85

QY 64 PMWIVDKRTETVAATGNIIRKIIHELGLCGHIEDIAVNSKYQGGLGKLLIDQLVTTIG 123  
Db 86 VTVEDVTLGQIVATATLIIHKFIHSCAKRGVEDVVSDCRGKQLKLLSTLTLLS 145

QY 124 FDYGVCIILDCDEKNVPEYKCGFS 149  
Db 146 KKLNCYKITLECLPQNVGFYKVGVT 171

RESULT 12  
AAG67123  
ID AAG67123 standard; protein; 184 AA.  
XX  
XX AAG67123;  
XX  
DT 13-NOV-2001 (first entry)  
XX  
XX Amino acid sequence of human 50287 transferase.  
XX  
XX Human; transferase; 25324 transferase; 50287 transferase; cancer;  
KW 28899 transferase; 47007 transferase; 42967 transferase; angiogenesis;  
KW kynurenine aminotransferase; phosphatidyl transferase; brain tumour;  
KW cell proliferation; cell differentiation.  
XX  
XX Homo sapiens.  
XX  
XX WO200164904-A2.  
PN  
XX  
XX 07-SEP-2001.  
PD  
XX  
XX 27-FEB-2001; 2001WO-US006463.  
PF  
XX  
XX 29-FEB-2000; 2000US-0185711P.  
PR  
XX  
XX (MILL-) MILLENNIUM PHARM INC.  
PA  
XX  
XX Meyers R, Macbeth KJ, Rudolph-Owen L;  
PI  
XX  
XX WPI; 2001-550186/61.  
DR N-PSDB; AAH75151.  
XX  
XX Novel transferase polypeptides and polynucleotides encoding the  
PT polypeptides, useful for the diagnosis or the identification of a  
PT compound capable of treating cancer or a disorder characterized by

PT aberrant angiogenesis.

PS Claim 4; Fig 4; 185pp; English.

XX The present sequence represents a human transferase polypeptide. The

CC specification describes 25324, 50287, 28899, 47007 and 42967

CC transferases. The 25324 transferase is homologous to kynurenine

CC aminotransferase. The 47007 transferase is homologous to phosphatidyl

CC transferase. The transferase polypeptides and polynucleotides may be used

CC for the diagnosis or the identification of a compound capable of treating

CC cancer (e.g., lung, breast, or colon) or a disorder characterised by

CC aberrant angiogenesis (e.g., brain tumour angiogenesis). Modulators of

CC the transferases may be used to treat the above, or modulate cellular

CC proliferation and/or differentiation, or a subject at risk of the above

CC disorders

XX

SQ Sequence 184 AA;

Query Match 24.9%; Score 208; DB 4; Length 184;

Best Local Similarity 32.9%; Pred. No. 3.1e-14;

Matches 48; Conservative 25; Mismatches 61; Indels 12; Gaps 2;

Qy 5 DGFYIRMEGDLQ-VTETLKVLTTVTITPESFKLIKYWNEATVWMDNEDKKIMQYN 63

Db 37 EGLVRLPLCTADLNRGPFKVLGQLTETGVSPQFMKSFEHM-----KKS GDY 85

Qy 64 PMVIVDKRTETVAATGNIIEERKIIHELGLCHIEDAVNSKYOGGLKLLIDQLVTIG 123

Db 86 VTVVEDVTLGQIVATATLIIIEHKFIHSCAKRVEDVWVSDECRGKLGKLLSTLTLLS 145

Qy 124 FDYGCYKIILDCDEKNVFKYKCGFS 149

Db 146 KKLNCYKITLECLPQNVGFKFGYT 171

RESULT 13

ADR10847

ID ADR10847 standard; protein; 184 AA.

AC ADR10847;

XX

XX 07-OCT-2004 (first entry)

XX Human FLJ10607, modifier of axin pathway.

XX Human; FLJ10607; axin; cytosstatic.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Domain 89..171

XX /note= "Acetyltransferase (GNAT) domain"

XX WO2004061086-A2.

XX 22-JUL-2004.

XX 29-DEC-2003; 2003WO-US041619.

XX 30-DEC-2002; 2002US-0436965P.

XX (EXEL-) EXELIXIS INC.

XX Gendreau SB, Dora EG, Lickteig K;

XX WPI; 2004-534377/51.

XX N-PSDB; ADR10844, ADR10846.

XX REFSEQ; XM\_085119.

XX Use of FLJ10607 polypeptide or nucleic acid for identifying a candidate

XX Axin pathway modulating agent, modulating an Axin pathway of a cell, or

XX diagnosing a disease, e.g. cancer in a patient.

XX

PS Disclosure; SEQ ID NO 4; 55pp; English.

XX The present sequence is that of a human FLJ10607 polypeptide. This is the

CC human orthologue of a Caenorhabditis elegans protein that has been shown

CC to have the ability to modify the axin pathway. The invention provides

CC methods for using FLJ10607 axin modifier genes and polypeptides to

CC identify FLJ10607-modulating agents that are candidate therapeutic agents

CC useful for the diagnosis and treatment of disorders associated with

CC defective or impaired axin function and/or FLJ10607 function, especially

CC proliferative disorders such as cancer. Preferred FLJ10607-modulating

CC agents specifically bind to FLJ10607 polypeptides and restore axin

CC function, or are nucleic acid modulators such as antisense oligomers and

CC RNA interference (RNAi) agents that repress FLJ10607 gene expression or

CC product activity, e.g. by binding to and inhibiting the respective

CC nucleic acid (i.e. DNA or mRNA). Candidate FLJ10607 modulating agents are

CC tested with an assay system comprising a FLJ10607 polypeptide or nucleic

CC acid.

XX

SQ Sequence 184 AA;

Query Match 24.9%; Score 208; DB 8; Length 184;

Best Local Similarity 32.9%; Pred. No. 3.1e-14;

Matches 48; Conservative 25; Mismatches 61; Indels 12; Gaps 2;

Qy 5 DGFYIRMEGDLQ-VTETLKVLTTVTITPESFKLIKYWNEATVWMDNEDKKIMQYN 63

Db 37 EGLVRLPLCTADLNRGPFKVLGQLTETGVSPQFMKSFEHM-----KKS GDY 85

Qy 64 PMVIVDKRTETVAATGNIIEERKIIHELGLCHIEDAVNSKYOGGLKLLIDQLVTIG 123

Db 86 VTVVEDVTLGQIVATATLIIIEHKFIHSCAKRVEDVWVSDECRGKLGKLLSTLTLLS 145

Qy 124 FDYGCYKIILDCDEKNVFKYKCGFS 149

Db 146 KKLNCYKITLECLPQNVGFKFGYT 171

RESULT 14

ADR10848

ID ADR10848 standard; protein; 184 AA.

AC ADR10848;

XX

XX 07-OCT-2004 (first entry)

XX Human FLJ10607, modifier of axin pathway.

XX Human; FLJ10607; axin; cytosstatic.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Domain 89..171

XX /note= "Acetyltransferase (GNAT) domain"

XX WO2004061086-A2.

XX 22-JUL-2004.

XX 29-DEC-2003; 2003WO-US041619.

XX 30-DEC-2002; 2002US-0436965P.

XX (EXEL-) EXELIXIS INC.

XX Gendreau SB, Dora EG, Lickteig K;

XX WPI; 2004-534377/51.

XX N-PSDB; ADR10845.

XX REFSEQ; NM\_198066.2.

XX Use of FLJ10607 polypeptide or nucleic acid for identifying a candidate

XX Axin pathway modulating agent, modulating an Axin pathway of a cell, or

XX



**THIS PAGE IS BLANK**

## Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

**When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.**

**THIS PAGE IS BLANK**

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☒ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**

**THIS PAGE IS BLANK**